

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 16:07:29 ; Search time 13988 Seconds

(without alignments)  
11511.316 Million cell updates/sec

Title: US-09-720-383C-9

Perfect score: 3936  
Sequence: 1 cttccctcctcgtcgtcgcg.....ctgagtcgttccttcgaa 3936

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045413386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genembl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: gb\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vit:\*  
38: em\_sy:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3935	100.0	3968	8	AF200531	AF200531 Zea mays
2	2543	64.6	3538	8	AF200530	AF200530 Zea mays
3	2241	56.9	3812	8	AF200532	AF200532 Zea mays
4	1531.4	38.9	3847	8	AF304374	AF304374 Nicotiana
5	1509	38.3	3255	6	AX507835	AX507835 Sequence
6	1489.8	37.9	3444	8	AF062485	AF062485 Arabidops
7	1434.2	36.4	3771	8	AF027173	AF027173 Arabidops
8	1431.8	36.4	3255	6	AX505864	AX505864 Sequence
9	1429.4	36.3	3828	6	AR267558	AR267558 Sequence
10	1429.4	36.3	3828	6	AX030944	AX030944 Sequence
11	1429.4	36.3	3828	6	BD022677	BD022677 Manipulat
12	1378.4	35.0	3725	8	AF200526	AF200526 Zea mays
13	1364	34.7	3752	8	AF200525	AF200525 Zea mays
14	1335.4	33.9	171376	8	AC104487	AC104487 Oryza sat
15	1332.4	33.9	3763	8	BT008654	BT008654 Arabidops
16	1330.8	33.8	3673	6	AR267560	AR267560 Sequence
17	1330.8	33.8	3673	6	AX030948	AX030948 Sequence
18	1330.8	33.8	3673	6	BD022679	BD022679 Manipulat
19	1326.6	33.7	3603	6	AR267557	AR267557 Sequence
20	1326.6	33.7	3603	6	AX030942	AX030942 Sequence
21	1326.6	33.7	3603	6	BD022676	BD022676 Manipulat
22	1321.4	33.6	3676	8	AF200529	AF200529 Zea mays
23	1301.8	33.1	3723	8	AF150630	AF150630 Gossypium
24	1298.2	33.0	3722	6	AX652952	AX652952 Sequence
25	1294.8	32.9	150778	2	AP004298	AP004298 Oryza sat
26	1289.4	32.8	3745	8	AF200528	AF200528 Zea mays
27	1279.4	32.5	3795	8	AF200533	AF200533 Zea mays
28	1279.4	32.5	3799	8	AX338680	AX338680 Sequence
29	1272.4	32.3	3729	8	BT002335	BT002335 Arabidops
30	1269.2	32.2	3614	6	AR267559	AR267559 Sequence
31	1269.2	32.2	3614	6	AX030946	AX030946 Sequence
32	1269.2	32.2	3614	6	BD022678	BD022678 Manipulat
33	1269.2	32.2	3682	8	AF027174	AF027174 Arabidops
34	1268.4	32.2	3264	6	AX653232	AX653232 Sequence
35	1250.6	31.8	3595	8	AF081534	AF081534 Populus x
36	1198	30.4	3081	8	AF088917	AF088917 Arabidops
37	1196.8	30.4	3355	8	AY139754	AY139754 Arabidops
38	1196.4	30.4	3081	8	BT004543	BT004543 Arabidops
39	1192.6	30.3	3277	8	AY095297	AY095297 Populus t
40	1192.4	30.3	3181	8	BT006111	BT006111 Arabidops
41	1192.4	30.3	3377	8	BT005710	BT005710 Arabidops
42	1184	29.8	3168	8	AF458083	AF458083 Arabidops
43	1171.2	29.7	2828	8	AF200527	AF200527 Zea mays
44	1170	29.7	2830	6	AX338676	AX338676 Sequence
45	1163.4	29.6	130300	8	AP005824	AP005824 Oryza sat

## ALIGNMENTS

RESULT 1  
AF200531  
LOCUS AF200531 3968 bp mRNA linear PLN 31-AUG-2000

DEFINITION Zea mays cellulose synthase-7 (Cesa-7) mRNA, complete cds.

ACCESSION AF200531  
VERSION AF200531.1 GI:9622885

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 3968)  
Holland, N., Holland, D., Helencjars, T., Dhugga, K.S.,  
Xocoostle-Cazares, B. and Delmer, D.P.

TITLE	A comparative analysis of the plant cellulose synthase (Cesa) gene family
JOURNAL	Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE	20398328
PUBMED	10938350
REFERENCE	2 (bases 1 to 3668)
AUTHORS	Dhugra, K.S. and Helentjaris, T.G.
TITLE	Direct Submission
JOURNAL	Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES	Location/Qualifiers

Db	1381	GGTTCCTTTCTGCAACGGTACAAATTTGAACTTCGCGCTCCAGGTGTACTTCCAA	1440
Qy	1442	GAAGAATGACTACTTTGAAAGCAAGGTGGCAGCAACTTTGTATGGAGAGAGAGCAAT	1501
Db	1441	GAAGATAGACTACTTTGAAAGCAAGGTGGCAGCAACTTTGTATGGAGAGAGAGCAAT	1500
Qy	1502	GAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATCCCTTATGTTGCCAAGCCAGAA	1561
Db	1501	GAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATCCCTTATGTTGCCAAGCCAGAA	1560
Qy	1562	AGTTCCTGAAGAAGATGAGCAATCCAGATGCGAAACCCCTGGCTGGAAACAATGTTCC	1621
Db	1561	AGTTCCTGAAGAAGATGAGCAATCCAGATGCGAAACCCCTGGCTGGAAACAATGTTCC	1620
Qy	1622	TGATCATCCTGGAAATGATTCAGGTCTTCCTCTGGCCAAAGCGAGGCTTGACTGAGGG	1681
Db	1621	TGATCATCCTGGAAATGATTCAGGTCTTCCTCTGGCCAAAGCGAGGCTTGACTGAGGG	1680
Qy	1682	AAATGAACCTGCCAGATTTGTTATGTTCTAGAGAGAAACGACCGCTATTAACAATCA	1741
Db	1681	AAATGAACCTGCCAGATTTGTTATGTTCTAGAGAGAAACGACCGCTATTAACAATCA	1740
Qy	1742	TAAAGAACTGGTGTATGAAATGCAATGGTCCGAGTCTCTGTGTATTAACAATGCTCC	1801
Db	1741	TAAAGAACTGGTGTATGAAATGCAATGGTCCGAGTCTCTGTGTATTAACAATGCTCC	1800
Qy	1802	ATATTTGTTAAACCTTGGATGTGATCAGTACATCAACAACGCAAGGCTATTAAGGAAGC	1861
Db	1801	ATATTTGTTAAACCTTGGATGTGATCAGTACATCAACAACGCAAGGCTATTAAGGAAGC	1860
Qy	1862	AATGTGTTTATGATGAGCCCTTTACTAGAAAGAGGTTTGTATGATACAGTTCCTCA	1921
Db	1861	AATGTGTTTATGATGAGCCCTTTACTAGAAAGAGGTTTGTATGATACAGTTCCTCA	1920
Qy	1922	AAGATTTGATGGGATTTGATCCGCATGACCGATATGCTAACCGGAATGTTGCTTTTGA	1981
Db	1921	AAGATTTGATGGGATTTGATCCGCATGACCGATATGCTAACCGGAATGTTGCTTTTGA	1980
Qy	1982	TATCAACATGAAGGTTTGGATGATTCAGGATCCAAATTTATATGTGATCATGATGCT	2041
Db	1981	TATCAACATGAAGGTTTGGATGATTCAGGATCCAAATTTATATGTGATCATGATGCT	2040
Qy	2042	ATTTAGAAAGCAGGCAATTATATGTTATGATGCCCCCAAAACAAGAACCCATCAAG	2101
Db	2041	ATTTAGAAAGCAGGCAATTATATGTTATGATGCCCCCAAAACAAGAACCCATCAAG	2100
Qy	2102	GACTTGCACATGCTGGCCCAAGTGGTCTTTGCTGTGCTTGGCAATAGAAACA	2161
Db	2101	GACTTGCACATGCTGGCCCAAGTGGTCTTTGCTGTGCTTGGCAATAGAAACA	2160
Qy	2162	AAAGAGACTACCAAAACCAAAACAAGAAAGAAAGTATATATTTTCAAGAAAGAGA	2221
Db	2161	AAAGAGACTACCAAAACCAAAACAAGAAAGAAAGTATATATTTTCAAGAAAGAGA	2220
Qy	2222	GAACCAATCCCTGCAATGCTCTTGGTGAATTTGACGAACCTGCTCAGGAGCTGGAA	2281
Db	2221	GAACCAATCCCTGCAATGCTCTTGGTGAATTTGACGAACCTGCTCAGGAGCTGGAA	2280
Qy	2282	TGAAGAGCCCGGTATTTGTAATCAACAAAATTGAAAAAGAAATTTGGCCATCTTCTGT	2341
Db	2281	TGAAGAGCCCGGTATTTGTAATCAACAAAATTGAAAAAGAAATTTGGCCATCTTCTGT	2340
Qy	2342	TTTGTATCATCCACATTCGAGAAATGGTGGAACTTTGAAGATGCAAGTCTGCTTC	2401
Db	2341	TTTGTATCATCCACATTCGAGAAATGGTGGAACTTTGAAGATGCAAGTCTGCTTC	2400
Qy	2402	TCTTTTGAAGAAGCTATACATGTCATAGTGTGGTTATGAACAACAAGACAGACTGGG	2461
Db	2401	TCTTTTGAAGAAGCTATACATGTCATAGTGTGGTTATGAACAACAAGACAGACTGGG	2460
Qy	2462	AAAAAGATTGGCTGATCTATGATCACTTACAGAAATATTTTAACTGGTTCAAGAT	2521

Db	2461	AAAGAGATTGGCTGGAATCTAATGATCAATGATACAGAAAGATATTTCTAACTGGTTCAAGAT	2520
Qy	2522	GCATTGTCATGGTTGGCGCGTCATATTATCTGCAATCTTAAACGGGTGCAATCAAGGTTTC	2581
Db	2521	GCATTGTCATGGTTGGCGCGTCATATTATCTGCAATCTTAAACGGGTGCAATCAAGGTTTC	2580
Qy	2582	TGCACCTCTGAATCTTTTCAGATCGTCTTACACAGGTGCTTGGTGGCTCTTGGGTCAT	2641
Db	2581	TGCACCTCTGAATCTTTTCAGATCGTCTTACACAGGTGCTTGGTGGCTCTTGGGTCAT	2640
Qy	2642	TGAGATCTTCTTCACAGCAATCAATGCCCTCTTGGTAAATGGGTAATGGTGGCGGCTCGAAT	2701
Db	2641	TGAGATCTTCTTCACAGCAATCAATGCCCTCTTGGTAAATGGTGGCGGCTCGAAT	2700
Qy	2702	TTTGGAAAGATTTTCTCAATCAATCAATCGATCGTATCTCTTGACATCTATATCCCTCTT	2761
Db	2701	TTTGGAAAGATTTTCTCAATCAATCAATCGATCGTATCTCTTGACATCTATATCCCTCTT	2760
Qy	2762	GGCTTACTGATCAATGCGCTGCGCATCTGTTAATGACAGGAAATTTATCACTCCAGAGCT	2821
Db	2761	GGCTTACTGATCAATGCGCTGCGCATCTGTTAATGACAGGAAATTTATCACTCCAGAGCT	2820
Qy	2822	GAATTAATGTCGCCAGCCCTGGGTGTCATGATCACTTTTATCTGCATTTTGTGTAAGACAT	2881
Db	2821	GAATTAATGTCGCCAGCCCTGGGTGTCATGATCACTTTTATCTGCATTTTGTGTAAGACAT	2880
Qy	2882	CCTGAATATGATGAGATGGATGGTGTGTTGAATTTGATGATCGTGGAGAGAAATGAGCACTGTG	2941
Db	2881	CCTGAATATGATGAGATGGATGGTGTGTTGAATTTGATGATCGTGGAGAGAAATGAGCACTGTG	2940
Qy	2942	GGTCAATGGAAGTGTGTCTCAACACTCTTGTGCTGTTCAGAGGACTTCTCAAGTCAT	3001
Db	2941	GGTCAATGGAAGTGTGTGTCTCAACACTCTTGTGCTGTTCAGAGGACTTCTCAAGTCAT	3000
Qy	3002	AGCTGGTGTATATCAAGCTTCAACCGTCGATCAATAAGGGTGGAGATGATGAGAGTTCTC	3061
Db	3001	AGCTGGTGTATATCAAGCTTCAACCGTCGATCAATAAGGGTGGAGATGATGAGAGTTCTC	3060
Qy	3062	AGAGCTATATATCAATTCATAATGGAATACCTTATATGATACCTCCACACTTGTGCTTAT	3121
Db	3061	AGAGCTATATATCAATTCATAATGGAATACCTTATATGATACCTCCACACTTGTGCTTAT	3120
Qy	3122	GAACCTCAATGGTGTGTGCTGACCTTTCATTAATGCGATCAATACCGATATGATGATG	3181
Db	3121	GAACCTCAATGGTGTGTGCTGACCTTTCATTAATGCGATCAATACCGATATGATGATG	3180
Qy	3182	GGGCCCCCTCTTTGGGAAGCTAATCTTTGCAATTTGGGTGATGTCCATCTTATCCCTT	3241
Db	3181	GGGCCCCCTCTTTGGGAAGCTAATCTTTGCAATTTGGGTGATGTCCATCTTATCCCTT	3240
Qy	3242	TCTCAAGGTTGGTGTGGAAGGCAAAACAGACACCAACGATTTGTATCGTCTGGTCCAT	3301
Db	3241	TCTCAAGGTTGGTGTGGAAGGCAAAACAGACACCAACGATTTGTATCGTCTGGTCCAT	3300
Qy	3302	TCTGCTGAGCTCAATCTTCTGCGCTCTTTGGGGTTCGGAATGATCCCTTCCCTGGGAAGA	3361
Db	3301	TCTGCTGAGCTCAATCTTCTGCGCTCTTTGGGGTTCGGAATGATCCCTTCCCTGGGAAGA	3360
Qy	3362	TGATGTGTCGCTTCTTGAAGAGTGTGGTTGATGTCGAATAGAGATGTCAGTGCATGAC	3421
Db	3361	TGATGTGTCGCTTCTTGAAGAGTGTGGTTGATGTCGAATAGAGATGTCAGTGCATGAC	3420
Qy	3422	TCCCCCAATCTGCAATATGCTTTGAAGTATTTTCTGCTGTTTGTCCCAATATCAAGTTC	3481
Db	3421	TCCCCCAATCTGCAATATGCTTTGAAGTATTTTCTGCTGTTTGTCCCAATATCAAGTTC	3480
Qy	3482	TGTGATTAAGACATGAATATGTCCTCAAGTTCTTTTGAATCAATGATGAACCTACTTAAT	3541
Db	3481	TGTGATTAAGACATGAATATGTCCTCAAGTTCTTTTGAATCAATGATGAACCTACTTAAT	3540
Qy	3542	ATCTGAAGATATATCTGGGGGAAATATGAGAGCTGCGGCAATCCTTGTGCAATTTGGGCGGT	3601
Db	3541	ATCTGAAGATATATCTGGGGGAAATATGAGAGCTGCGGCAATCCTTGTGCAATTTGGGCGGT	3600

Oy		3602	GGAATACAGCATATCAAGTGGTTGATTGTGCAGAACTCTTAATACTTGTCGCAATAT	3661
Db		3601	GGAAATACAGCATATGCAAGTGGTTGATTGTGCAGAACTCTTAATACTTGTCGCAATAT	3660
Oy		3662	AGATGGCGTCCAGCCGAAACGCAAGGATTTTGATTCTGCACTGCTCCGCTGACAACCTT	3721
Db		3661	AGATGGCGTCCAGCCGAAACGCAAGGATTTTGATTCTGCACTGCTCCGCTGACAACCTT	3720
Oy		3782	TTTATGTATGACCTGTTTCATTTGGAGGGCTGTTGTCATTACATGTTGCTATATCTAAGAAAA	3841
Db		3781	TTTATGTATGACCTGTTTCATTTGGAGGGCTGTTGTCATTACATGTTGCTATATCTAAGAAAA	3840
Oy		3842	ACAGAAATATTAGCATTAACTATATAGTTAATAATTAAGTATGAATAGCGCTGTTTTTGT	3901
Db		3841	ACAGAAATATTAGCATTAACTATATAGTTAATAATTAAGTATGAATAGCGCTGTTTTTGT	3900
Oy		3902	GTTGTCCTGTAATCATCTGAGTTGGTTTGTGAAA	3936
Db		3901	GTTGTCCTGTAATCATCTGAGTTGGTTTGTGAAA	3935
RESULT 2				
AF200530		3558 bp	mRNA	PLN 31-AUG-2000
LOCUS				
DEFINITION	Zea mays cellulose synthase-6 (Cesa-6)		mRNA, complete cds.	
ACCESSION	AF200530			
VERSION	AF200530.1		GI:9622883	
KEYWORDS				
SOURCE				
ORGANISM	Zea mays			
	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
	clade; Panicoidae; Andropogoneae; Zea.			
	1 (bases 1 to 3538)			
	Holland,N., Holland,D., Helentjaris,T., Dhugra,K.S.,			
	Xocostele-Cazares,B. and Delmer,D.P.			
	A comparative analysis of the plant cellulose synthase (Cesa) gene			
	family			
TITLE	Plant Physiol. 123 (4), 1313-1324 (2000)			
JOURNAL				
MEDLINE	203983328			
PUBMED	109383350			
REFERENCE	2 (bases 1 to 3538)			
AUTHORS	Dhugra,K.S. and Helentjaris,T.G.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA			
FEATURES				
SOURCE	Location/Qualifiers			
	1..3538			
	/organism="Zea mays"			
	/mol_type="mRNA"			
	/db_xref="taxon:4577"			
	/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc."			
gene	1..3538			
	/gene="Cesa-6"			
CDS	43..3222			
	/gene="Cesa-6"			
	/note="CESA-6"			
	/codon_start=1			
	/product="cellulose synthase-6"			
	/protein_id="AAF89966.1"			
	/translation="MDNRNGVCQICGPDVGNRPDGPBPVACNBAPICRDCYEVSRBGRNCPCQCKTRPKRKFGKCAVPDSDBDGYDLLENFMSDHSOYLASMLAHMSYSGADLDGVDPFHPIPNYPILITNGMVDLIPFDHALVPSVGGGGRKLPLPYADPNLPVOPRSMDPSKDIAAYGYSVANKERESWKQGEHMHQIRNDGGDDGDADDLPLDMDEQPLSRKILPSSQINPRMIITRLVYLCPFFHYRWHPVDFALMLLS			

BASE COUNT	876 a	813 c	928 g	920 t	1 others
ORIGIN					
Query Match	64.6%;	Score 2543;	DB 8;	Length 3538;	
Best Local Similarity	86.9%;	Pred. No. 0;			
Matches 2821;	Conservative	0;	Mismatches 415;	Indels 9;	Gaps 2;
QY	192	CTCGTCGTCATCCGCGCGCAGCGCGATCCGCGCCGAGCGCGCGGAGCAGACGGG	251		
DB	1	CTCGTCGTCATCCGCGCGCAGTGGCGAGCCAGGGCCGAGCCCATGAGCACGCGGAA	CGGC	60	
QY	252	CAGGTGTGCCAGATTTCCGGCGACGACGTGGGCTTTGCCCCCGCGGGGACCCCTTG	CGTG	311	
DB	61	CAGGTGTGCCAGATTTTCCGGCGACGACGTGGGGGGGCAACCCCGACGGGAGCC	CTTGCGTG	120	
QY	312	GCCTGTGCAACGAGTGCGGCTTCCCCGCTGCGCGGGACTGCTACGAAATACGAGCGCGGGAG	371		
DB	121	GCCCTGCAACGAGTGCGGCTTCCCCATTTGCCCGGGACTGCTACGAGTACGAGCGCGCGAG	180		
QY	372	GGCAGCGCAGACCTGCCCCCACTGTCAGACCTCGATACAGCGCCTTCAAGGGCTGCGAA	CGT	431	
DB	181	GGCAGCGCAGACCTGCCCCCACTGTCAGACCTCGATTAAGGGCTTCAAGGGGTGCGCGGC	240		
QY	432	GTCACCGGTGACGAGAGGAGAGACGGCGTGTGATGACTGACCTGACACGAGTTCAATTGC	--	488	
DB	241	GTCGCCGGGGACGAGGAGGAGAGACGGCGTGTGACCACTGAGGAGACGAGTTCACTGAGAC	300		
QY	489	GAGCGGCATGACTGCGAGTCTGTGGCGCGAGTCCATGTCTACCGGCCCATGAGCTTACGGC	548		
DB	301	GACACGACGACCTTCCCACTGCTCGCGGAGTCCATGTCTCACCGCCCATGAGCTTACGGC	360		
QY	549	CGTGGAGGTGACCTTAATGGCGCGCCACAGCTTTCCAGTCAACCCCAATGTTTCACTC	608		
DB	361	CGCGGCCCCGACCTTCGACGGCGTGGCGGACGATTTCAACCCCAATGTTTCCCTTC	420		
QY	609	CTCACCAACGGGCAATGTGTGATGACATCCACCGAGCAGACGCGCTGTGTCTTCT	668		
DB	421	CTCACCAACGGGACAGATGTGTGATGACATCCCGCGGACAGCACGCGCTTGTCTCCCTCG	480		
QY	669	TTTCATGGGTGTGGGGGAAAGAGATATATCCCTTCTTATGTGGATATCCAGCTTACT	728		
DB	481	TTTCGTGGGTGTGGGGGAAAGAGATATATCCCTTCTTCCGTATCGGGATATCCCACTTCT	540		
QY	729	GTCGACCCAGGTGTATGACCCATCCAGAGATCTTGTCTGATATGGGTATGTATGTT	788		
DB	541	GTCGACCCAGGTGTATGACCTTCCAGAGATTTCCGCGGATTTGGCTACGGGAGCTTA	600		
QY	789	GCTTGAAGGACCGATGTAGAAATTGGAGCAGACAGAGAGAGATGACACGAGCGGG	848		
DB	601	GCAITGAAGGAGAGATGTAGAGCTGTGAAGCAGAAAGAGAGAGATGACACGAGCGAG	660		
QY	849	AATATAGTGTGTGTGATGTATGTGACGATGCTGATTTACCTATATGATGAAACGAGA	908		
DB	661	AACATATGCGCGCGCGATGTATGTGATGACGATCTACCACTTAATGATGAAAGCTTGA	720		
QY	909	CAACAACTGTCCAGAGAAATTTCACTTCATCCAGACGAGATTAATCAATATAGATGATT	968		
DB	721	CAGCAATTGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTATAGATGATT	780		



QY	965	ATCATTTATTCGGCTGTGGTTTGGGGTCTTCTTCCATCACTCCAGATGATGATCCGGTG	1028
Db	781	ATAAATAATTCGGCTAGGTGGTTTGGTTTCTTCTTCCATCACTCCAGATGATGATCCGGTG	840
QY	1029	AATGATGACATTTGCTTTTGCGCTCAATATCTGTATCTGTGAATATCGGTTTGGCATGCT	1088
Db	841	CCTGATGACATTTGCTTTTATGCTCATATCTGTGATCTGTGAATTTGGTTTGGCATGCT	900
QY	1089	TGGATTTCTTGATCAATTCGCCAAGGTGTTCCCTATTGAGAGAGACTTAACCTAGACCGG	1148
Db	901	TGGATTTCTTGACCAAGTTTCCAAAGTGTTCCTATTCAGAGGGAAACCTATCTTGACCGG	960
QY	1149	CTGTCACTGAGGTTTGGACAGAGAAGCCAGCCACTCAACTTGTCTCAATTGATTTCTTT	1208
Db	961	CTGAGTTTAAAGGTTTGAACAAGAAAGGGCACTCTTCTCAACTCGCCCTGTGTAATTTCTTT	1020
QY	1209	GTCAGTACGGTGTATCCCTTAAAGGAACCTCCTTTGGTGCACAACAATACTGTTCTATCT	1268
Db	1021	GTCAGTACGGTGTGATCCCTTGAAAGGAACCTCCATTGGTGCACCTGCTAAATAGTGTCTATCT	1080
QY	1269	ATCCTTTCCGTGGAATTAATCTGTTGATTAAGTTTCTGTATGTTTCTGATGATGAGTGCCT	1328
Db	1081	ATCCTTTCCGTGGAATTAATCAAGTTGATTAAGGTTTCACTGTAAGTTTCTGATGATGAGTGCCT	1140
QY	1329	GCAATGCTTAACCTTTGAAGCAATTATCTGAACAATCTGAATTTGCAAGAAATGGGTTCT	1388
Db	1141	GCCATGCTTGAACCTTTGAAGCAATTATCTGTAACAATCTGAATTTGCAAGAAATGGGTTCT	1200
QY	1389	TTCTGCAAAAGGTATGCAATTTGGAACCTCGCGCTCGAGTGTGATCTTCCACAGAAAGTA	1448
Db	1201	TTCTGCAAAAGGTATGCAATTTGGAACCTCGCGCTCGAGTGTGATCTTCCACAGAAAGTA	1260
QY	1449	GACTACTTGAAGACAGAGGTGGCAGACAAACTTTGTTAAGGAGAGAGACATGAAGAGA	1508
Db	1261	GACTACTTGAAGACAGAGGTGGCAGACAAACTTTGTTAAGAGAGAGACATGAAGAGA	1320
QY	1509	GAGTATGAGGAATTAAGGTGAGATCAATGCTTTAGTGTCCAAAGCCCAAGAAAGTCTCT	1568
Db	1321	GAGTATGAGGAATTAAGGTGAGATCAATGCTTTAGTGTCCAAAGCCCAAGAAAGTCTCT	1380
QY	1569	GAAGAGAGTGGACAAATGCAAGATGGAACCCCTGCGCTGGAAACAAGTTCGTGATCAT	1628
Db	1381	GAAGAGAGTGGACAAATGCAAGATGGAACCTCCATGCGCCGGAAATAAAGTTCGTGATCAT	1440
QY	1629	CCTGGAATGATTCAGAGTCTTCTTGGCCAAAGCGGAGCCTTGACTGTGAGGGAAATGAA	1688
Db	1441	CCTGGAATGATTCAGAGTCTTCTTGGTCAAAAGTGTGGCCATGATGTGAAAGAAATGAG	1500
QY	1689	CTGCACGATTTGTTATGTTTCTAGAGAGAAACGACCAAGGCTATTAACCATCATAGAAA	1748
Db	1501	CTGCCTCATTTGTTATGTTTCTAGAGAGAAACGCGCACAGGCTATTAACCATCATAGAA	1560
QY	1749	GCTGTGCTATGATGATGATGTTGGTCCGAGTCTCTGTGTTACTAACAATGCTTCATATTG	1808
Db	1561	GCTGTGCTATGATGATGATGTTGGTCCGAGTCTCTGTGTTACTAACAATGCTTCATATTG	1620
QY	1809	TTAACCTTGAATGTGATCAGTCACTCAACAACAGCAGGCTATTAAGAGAGCAATGTGT	1868
Db	1621	CTGAACCTTGAATGTGATCAGTCACTAATCAATATATAGGCTATTAAGAGAGCAATGTGT	1680
QY	1869	TTTATGATGACCCCTTTACTAGAGAAAGGTTTGCTATGTATCAAGTTCCTCAAGAATTT	1928
Db	1681	TTTATGATGATGATCCTTTGCTTGAGAAAGAAAGTTTGCTATGTGCAAGTTCCTCAAGAATTT	1740
QY	1929	GATGGGATTTGATCGCCATGACCGGATATGCTAAACCGGAATGTTGCTTTTGTGATATCAAC	1988
Db	1741	GATGGGATTTGATCGCCATGATGATGATGCTAAACGAATAATGTTGCTTTTGTGATATCAAC	1800
QY	1989	ATGAAGAAGTTTGGATGATGATTCAGGGGTCCAATTATGTTGGATCGATGTGTATTTAGA	2048
Db	1801	ATGAAGAAGTTTGGATGATGATTCAGGGGCCCAATTATGTTGGGATCTGAGATGTGTCTTCA	1860
QY	2049	AGGACGACATTAATGTTATGATGATCCGCCCAAAACAAAGAGCCATCAGGACTGTGC	2108

Db	1861	AGGAGGAGATTATATGCGTACGAGTGCCTCCAAAA	CAAGGAGCCACCTCAAGAACTTGG	1920
Oy	2109	AAGCTGTGGCCCAAGTGGTGTCTTTTGCTGTGCTGCTTGGCAATAGAGACAAAGAG	2166	
Db	1921	AACGTGTGGCCAAAGTGGTGCATTTGCTGTGCTGTGTTTGGTAA	CAGAAACCAAGAG	1980
Oy	2169	ACTACCAAAACCCAAAACAGAGA-----AGAAAAAGTATATTTTTC	CAAGAAAGAG	2222
Db	1981	AAGACCAAGACCTCTAAACCTAAATTTGAGAAAGATTAAGAA	ACTTTTAAAGAAAGGAA	2040
Oy	2223	AACCAATCCCTGCATATGCTCTTGGTGAATTTGACGAAGTCTCCAGAGCTGAGAT	2288	
Db	2041	AATCAAGCCCCGTGATATGCTCTTGGTGAATTTGATGAAGCCGCTCCAGAGCTGAAAT	2100	
Oy	2283	GAAAGGCGGGTATTTGTAATCAACAAAAATTAAGAAAGAAATTTGGCAATCTTCTGT	2342	
Db	2101	GAAAGGCTAGTATTTGTAATCAACGAAGTTGAAAGAAATTTGGCAAGCTTCACT	2166	
Oy	2343	TTTGTATCATCCACACTTCTCGAGAAATGTGGAACCTTGAAGATGCAAGTCCGTCT	2402	
Db	2161	TTTGTGCATCCACACTTCTTGAAGAAATGTGGAAGCCCTGAAGATGTGCAAGCTTCT	2220	
Oy	2403	CTTTTGAAGAAGCTATACATGTATTAAGTTGTGTTATGAAACAACAAGACTGGGA	2462	
Db	2221	CTTCTGAAGGAGCTATACATGTATCATGTTGTGATATGAAAGCAAAACAGGCTGGGA	2280	
Oy	2463	AAAGAGATTGGCTGAGATCTATGATTCATGATACAAAGATTTCTAATCTGGTTCAAGAT	2522	
Db	2281	AAAGATATTTGGTATGATTTATGATTCATGATACAAAGATTTCTAATCTGGTTAAGAT	2340	
Oy	2523	CATTGTATGATGGTGGCGGTCAATTTACTGTGATACCTTAAACGGGTTGATTCAAAGTCT	2588	
Db	2341	CATGTCATGGTGGCGGTCAATTTACTGTGATACCTTAAACGGGCTTCAAGTCTC	2400	
Oy	2583	GCACCTTGAATCTTTCAAGATCGTCTTCAACAGGTGCTTGGTGGGCTCTTGGTCTAT	2642	
Db	2401	GCACCTTGAATCTTTCCGATCGTCTTCAACAGGTGCTTGGTGGGCTCTTGGTCTAT	2460	
Oy	2643	GAGATCTTCTTCAGCAATCAATGGCCCTTTTGGATGGGATATGGTGGCGGTCTGAATTT	2702	
Db	2461	GAAATTTTGTTCAGCAACCTGCGCTCTGTGATGGGATATGGTGGGCTGAATTTT	2520	
Oy	2703	TTGGAAGATTTTCTCATCAATCCATCGATCGTGTATCTTGGACATCTATCCCTCTG	2762	
Db	2521	CTGGAAGGTTTTCATCAATTAATCTCATATGTATACCTTGGACATCTATCCCTCTG	2580	
Oy	2763	GCTTACTGTACATTTGCGGCATCTGTTTATATGACAGGGAATTTATCACTCAGAGCTG	2822	
Db	2581	GCTTACTGTACATTTGCGGCATCTGTTTATATGACAGGGAATTTATCACTCAGAGCTT	2640	
Oy	2823	AATTAATGTGCGACGCTGTGGTTCATGTCACTTTTATCTGCATTTTGTCTACAGACAT	2882	
Db	2641	AACATATTTGCGACGCTGTGGTTCATGTCACTTTTATCTGCATTTTGTCTACAGACAT	2700	
Oy	2883	CTAGAAATGAGATGAGATGGTGTGGAATGATGACGATGGAGGAATGACATTCG	2942	
Db	2701	CTAGAAATGAGATGAGATGGTGTGGAATGATGACGATGGAGGAATGACATTCG	2760	
Oy	2943	GTCATTTGAGAGTGTGTCTCACACTTCTTGTCTGTCTTCCAGGACCTTCTCAAGTATA	3002	
Db	2761	GTCATTTGAGAGGCGTGTCTCACACTTCTTGTCTGTCTTCCAGGACCTTCTCAAGTATA	2820	
Oy	3003	GCTGTGTGTATACAGACTTCAACGTGACATCAAGGCTGAGATGATGAGAGTTCTCA	3062	
Db	2821	GCTGTGTGTATACAGACTTCAACGTGACATCAAGGCTGAGATGATGAGAGTTCTCA	2880	
Oy	3063	GAGCTATATACATTTCAAGTGAAGTATTTGATATCTCTCAACCTTGTCTTATTTG	3122	
Db	2881	GAGCTATATACATTTCAAGTGAAGTATTTGATATCTCTCAACCTTGTCTTATTTG	2940	
Oy	3123	AACCTCAATGTGTGTGCTGGCGTTTCAATATGCAATCAATTAAGATATGATCATG	3182	



Db	964	GGCGACGATGCAGATCTGGCACTAATGATGAAGCTAGGACGCCATTGTCAGAAAAGCTC	10233
Qy	930	CCACTTCATCAAGCCAGATTATCATATAGATGATATATCAATAATTCGGCTGTGGTT	989
Db	1024	CCATATATCATCAAGCCGAATTATATCCCTACAGATGATTTATTCGTTATCCGGTGGTGTT	10833
Qy	990	TTGGGGTTCTTCTTCCACTACCCGAGTATGATCCGGTGAATGATGCAATTTGCTTTGTGG	10449
Db	1084	TTGGGTTTCTTCTTCCACTACCGAGTATGATCCGGGGAAGATGATTCATTTGCAATTTGGG	11433
Qy	1090	CTCATATCTGTTATCTGTGAAATCTGGTTTGGCATGTCTTGGATTCTTGATTCATTTCCCA	11099
Db	1144	CTCATATCTGTTATCTGTGAAATCTGGTTTGGAGTGTCTGGATTCTTGATTCAGTTCCCA	12033
Qy	1110	AAGTGGTTCCCAATTGAGAGAGACTTAACCTAGACCCGAGCTGTCACTAGAGGTTCGACAAG	1169
Db	1204	AAGTGGCTTCCAAATGAGAGAGAGACTTACCTGGACCGTTGTCTACTAAGGTTTGAACAG	12633
Qy	1170	GAAGGCCACGCATCTCAACTTGTCTCAATTGAATTTCTTTGTCTAGTACGGTTGATCCCTTA	1229
Db	1264	GAAGGTCAACCCCTCAGCTTGGCTCAATACGACTTTTGTCTAGTACGGTTGATCCCA	13233
Qy	1230	AAGGAACCTCTTGTGTGACAAACAATATCTGTCTATCTATCTCTTCCGTGGATTATCT	1289
Db	1324	AAGGAACCTCTCTTGTGTGACACGCGAACACTGCTCTTCCATCTCTTCTGTGGAATTAATCCG	13833
Qy	1290	GTTGATTAAGTTTCTTGTGATCTTCTATGATGTGTGTGCTGCAATGCTAAAGCTTTGAAGCA	1349
Db	1384	GTTGAGAAAGTCTCCTGCTATGTCTTGTATGATGTGTGTGCTGCAATGCTTAAGTTTGAAGCA	14433
Qy	1350	TTATCTGAACAATCTGAAATTTGCAAGAATATGGTTTCTTTCTGCAACGGTACATATTT	1409
Db	1444	TTGTCTGAACAATCTGAAATTTGCAAGAATATGGGTTTCTTTCAAGAAAAGTTTAATATC	15033
Qy	1410	GAACCTGGCGCTCCAGAGTGTGATCTTCCAACAAGATAGACTACTTGAAGAACAAGTG	1469
Db	1504	GAGCTCGTGTCTCTGAGTGTGATCTTCCAACAAGATAGACTACTTGAAGAACAAGTT	15633
Qy	1470	GCAGCAAACTTGTGTAAGGAGAGAGACAAATGAAGAAGTATGAAGAAATTCAGAGTG	1529
Db	1564	GCTGCTTCAATTTGTGTAAGGAGAGAGAGGAGATGAAGAATACGAAGAAATTCAGAGTA	16233
Qy	1530	AGAAATCAATGCTTAGTATGTTGCCAAGACCCAGAAAGTTCCTGAGAAGATGACATGCA	1589
Db	1624	AGAAATCAATGCTTAGTATGTTGCCAAGAACCCAAAGGTTCTGAGGAAGATGACATGCA	16833
Qy	1590	GATGGAACCCCTGTGGCTGTGAACAACAATGTCGATCTCTGGAATGATTCAGGTATTC	17433
Db	1684	GATGGAACCCCTGTGGCTGTGAACAACAATGTCGATCTCTGGAATGATTCAGGTATTC	17433
Qy	1650	CTTGGCCAAAGGAGAGGCTGTGACCTGTAGAGGAAATGAACTGSCACCATTTGGTTATGTT	1709
Db	1744	CTTGGCCAAAGGAGAGGCTGTGTGTATGTGAAGAAATGAGTGTCTCGCTGTGTTATGTC	18033
Qy	1710	TCTAGAGAAACAACAAGCTATATTAACATCATTAAGAAGCTGTGTATGAATGCAATG	1769
Db	1804	TCTAGAGAAACAAGAGCCAGGTTTATACATCATCAAGAAGGCTGTGTGCAATGAATGCACTG	18633
Qy	1770	GTCGAGTCTCTGTCTGTACTAACAATGCTCCATATTTGTTAACTTGGATTGTGATCAC	1829
Db	1864	GTCGAGTCTCTGTGTCTGTACTAACAATGCTCATCTCAATTAAGAACTTGGACTGTGATCAC	19233
Qy	1830	TACATCAACAAGAGAGGCTATTAAGGAACAAATGCTTTTATGATGGAACCTTTACTA	1889
Db	1924	TACATCAACAAGAGAGGCTATTAAGGAACAAATGCTTTTATGATGGAATCCTTGGTG	19833
Qy	1890	GGAAAAGAGTTTGCTATGTACAGTTCCTCCAAAGATTTGATGGGATGTATGCTGCAATAC	1949
Db	1984	GGAAAAGAAAGTGTGTATGTACAGTTCCTCCAGAGGTTTGTATGATTTGAACAAAATGAT	20433
Qy	1950	CGATATGCTTAACCGGAATGTGTCTTTTGTGATTAACAATGAAGCTTTGGATGATAT	2009
Db	2044	CGATATGCTTAACCGGAATGTGTCTTTTGTGATTAACAATGAAGGTTTGGATGATAT	21033

QY	2010	CAGGCTCCAAATTATAGTTGGTACTGAGTGTGTATTTAGAAAGCAGGCAATTATATGTAT	2065
Db	2104	CAAGGACCATTATATAGTGGTACTGATATGTATTTTCAGACGGCAGGCACTGTATGTAT	2163
QY	2070	GATGCCCCCAAAACAAGAACCCACATCAAGGACTTGCACTGTGGCCCCAATGGTGC	2129
Db	2164	GATGCTCTTAAAGAAAGAACCCACATCAAGAACTTGCAACTGTGGCCCCAATGGTGC	2223
QY	2130	TTTTGCTGTGGCTCTTTGGCAATAGGAAGCAAAAGAACTACCAAAACCMAAACAG	2189
Db	2224	CTCTCTTGCTGTCTCAGAGAGAACATATATATATATAGATCTAATAACCAAGACGAG	2283
QY	2190	AAGAAAAATTATTTTTCAGAAAGAAAGAACCAATCCCTGATATGCTTTGGT	2249
Db	2284	AAGAGAAAAAGATTATTTTTCAGAAAGAAAGAACCAATCTCTGATATATGCTTTGGT	2343
QY	2250	GAAATTGACGAGCTGTCTCCAGAGCTGAGATATGAAAGCCGGTATTTGTAATCAACA	2309
Db	2344	GAAATATGATGAAGGTGTCTCCAGGCTGATATCAGAAAGCCGGAAATCTGTAATCAAC	2403
QY	22310	AAATTAGAAAAAATTTTGGCCATCTTCTGTTTTTTGTATCACTCAACTTCTGCAAA	2369
Db	2404	AAACTAGAGAAAGAAATTTGGGCACTCTTCTGTTTTTGTGCGATTAACATTTCTGAGAAC	2463
QY	2370	GGTGGAACTTTGAAGAGTGCAGTCTGCTTCTCTTTGAAAGAGTATATCATGTCA	2429
Db	2464	GGAGGGAACCTGGAAGAGGCGAACGTCCAGCTTCTCTTGAAGGAAGTATATCATGTTATC	2523
QY	2430	AGTTGTGTTATGAAACAAAGACAGACTGGGGAAAAAGATTTGGCTGATCTATGATCA	2489
Db	2524	AGCTGCGCTACGAAGCAAGACCGACGTGGGAAAAAGATTTGCTGATTTAGGATCG	2583
QY	2490	GTTACAGAAATATTTCACTGTTTCAAGTGCATTTGATGGTGGCGGTCAATTAC	2549
Db	2584	ATCACAGAGATATCTTGACTGATTTATAGATGACCTGCATGGCTGGCGGTCTATTTAC	2643
QY	2550	TGCATTACTTAAACGGTGTGCAATTCAAAGTTCTGCACTCTGATCTTTCAGATCGTCTT	2609
Db	2644	TGCATCCCGAAGCGGCGCTGCAATTAAGGTTCTGCGGCTCTGAACCTTTCGACCGTCTT	2703
QY	2610	CACCAAGTGTCTGGTGGGCTCTTGGGTCTATTTGAGATCTTCTTCAGCAATCATTTGCCCT	2669
Db	2704	CACCAAGTGTCTGGTGGGCTCTTGGGTCTATTTGAGATCTTCTTCAGCAATCATTTGCCCA	2763
QY	2670	CTTTTGGTATGGGTATGTGTGGGTCTGTAATTTTGGAAAGATTTTCTATCAATCACTCC	2729
Db	2764	CTTTTGGTATGGGTATGTGTGGGTCTGTAATTTTCTGAAAGGTTTCTTATATCAACTCC	2823
QY	2730	ATCGTGTATCTTGGACATCTATCCCTCTTGCTTACTGATATGCTCGCATCTGT	2789
Db	2824	ATCGTGTATCTTGGACATCTATCCCTCTTGCTTACTGATATGCTCGCATCTGC	2883
QY	2790	TTATTTGACAGGAAATTTTATCACTCCAGACTGAATATGTGCAACTGTGGTTATG	2849
Db	2884	CTGCTCAACGGGAAATTTTATCACTCCAGACTGAATATGCAATGTGCGCATGTGTTATG	2943
QY	2850	TCACTTTTATCTGCATTTTTCGTAACAGCATCTTGAATATGATGAGTGTGTGGA	2909
Db	2944	GCACTTTTATCTGCATCTCCGTGACCGGATCTCTGAAATATGAGTGGAGTGTGGCC	3003
QY	2910	ATTATATACGTGTGAGGAAATGAGCAGTTCTTGGGTCAATGGAAGTGTCTTCAACCTC	2969
Db	3004	ATGACACATGTGTGAGGAAATGAGCAGTTCTTGGGTCAATGGAAGCGTTTGGCGCATCTG	3063
QY	2970	TTTTGCTGTGTCAGGGAATTTCTCAAGCTATATGCTGTGTATATCAACCTTCAACGTG	3029
Db	3064	TTTGCGGTGTTTCCAGGGCGCTGTGAAAGTGTGTGCGGACATGACAGAACTTCAACGTG	3123
QY	3030	ACATCAAAAGGTGAGATGATGAGAGTCTTCAAGCTATATATCAATTCAATGAGCTAAC	3089
Db	3124	ACGTGGAAGCGCGGAGACGACGAGAGTCTTCCGAGCTGTATACCGTTCAAGTGAACAC	3183

QY 3090 TTAATGATACCTCTTACCTCTTCTTATGAACTTCAATGGTGTGCTGCTGCTGCT 3149  
 DB 3184 CTGCTATATACCCCGACGACGCTCTCTCTCTGAACTTCAATGGGGGTGGCGGGATC 3243  
 QY 3150 TCAATGCGATCAATAACGATATAGTCAATGGGGGCCCCCTCTTGGGAAGCTATCTT 3209  
 DB 3244 TCGAAGCGCATCAACAGGGTACGACTCTGGGGCCCCCTGTTGGGAAGCTCTTCTTC 3303  
 QY 3210 GCATTTGGGTGATGTGCTCACTTTTATCCCTTCTCAAGAGTTTGGTGGAGCAAAAC 3269  
 DB 3304 GCCTTCTGGGTGATGTGCTCACTGTACCCGTTCTCTCAAGAGTTGTTGGAGGCGAAG 3363  
 QY 3270 AGGACACCAACGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3329  
 DB 3364 AGGACCGCCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3423  
 QY 3330 TGGGTTGCGATGATGCTTCTTCTGCGAAGATGATGCTGCTGCTGCTGCTGCTGCT 3389  
 DB 3424 TGGGTTGCGGCTGCAACCGCTTCTGCGAAGGACGAGCGCGCTCTGAGAGAGTGTGC 3483  
 QY 3390 TTGATTTGCAACT 3402  
 DB 3484 CTGACTGCAACT 3496

RESULT 4  
 AF304374 3847 bp mRNA linear PLN 02-MAY-2001  
 LOCUS AF304374  
 DEFINITION Nicotiana alata cellulose synthase catalytic subunit (CesA1) mRNA.  
 ACCESSION AF304374  
 VERSION AF304374  
 KEYWORDS complete cds.  
 SOURCE AF304374.1 GI:13925880  
 ORGANISM Nicotiana alata (Persian tobacco)  
 Nicotiana alata  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.  
 REFERENCE 1 (bases 1 to 3847)  
 AUTHORS Doblin,M.S., De Melis,L., Newbiglin,E., Bacic,A. and Read,S.M.  
 TITLE Pollen tubes of Nicotiana alata express two genes from different beta-glucan synthase families  
 JOURNAL Plant Physiol. 125 (4), 2040-2052 (2001)  
 MEDLINE 21196092  
 PUBMED 11299383  
 REFERENCE 2 (bases 1 to 3847)  
 AUTHORS Doblin,M.S., De Melis,L., Newbiglin,E., Bacic,A. and Read,S.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-SEP-2000) Plant Biology, University of California, 1 Shields Ave, Davis, CA 95616, USA  
 FEATURES  
 source  
 1. 3847  
 /organism="Nicotiana alata"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4087"  
 /issue\_type="plastid"  
 /note="Genotype: S252"  
 1. 3847  
 /gene="CesA1"  
 103. 3378  
 /gene="CesA1"  
 /codon\_start=1  
 /product="cellulose synthase catalytic subunit"  
 /protein\_id="AAK49454.1"  
 /db\_xref="GI:13925881"  
 /translation="WDTKRLVAGSHNRNFFVINADVGRVTSVKSQICQICGD  
 EIEVYDGEFPICNECAPVRCQCYEYERREGNACPCQKTRFKIKGSPRVGDDE  
 DDEFDLIDHEFDYHGNPRVSEAFSSRLRGTHNNSGLTTSSEVDPALNIEPL  
 TYQEDDTTISADKHALIIPFMGKGVHPVSDSMSLPKPMWPKDLAVYGYGV  
 AKERMEDEKKKNDKLOVYKGGGKGGGDELDPDLPKMBGQPLSRLLPSS  
 SRLSPYRLIVLAVGLFPHYRITHPNVDALMLISIKIMPAWSVIRPOPRW  
 PPIVRETYIDRLIRYKEKPKGLADIDIFSTVDLKKPPLITATVLSIAVDIP  
 EDVSCVSDGAAMLTFEALSETSEPARRWVPCKKKNIEPRAPEWYFSQKVDYLNK

BASE COUNT 1061 a 719 c 925 g 1142 t  
 ORIGIN  
 Query Match 38.9%; Score 1531.4; DB 8; Length 3847;  
 Best Local Similarity 68.2%; Pred. No. 8.8e-286;  
 Matches 2252; Conservative 0; Mismatches 996; Indels 53; Gaps 7;  
 DE"

QY 139 AAGCGATGAGGCGGACGCGCGGCTGCTGCGGCTCCCAACCCGACGACTGCTG 198  
 DB 98 AAGTATGATGATTAAGGAGGCTAGTGTGCTGTTCAACAATAGGAATGAGTTGTG 157  
 QY 199 TCATCCGCGGACGCGGATCCCGGCGGACGCGCGGAGCAGAGCGGCGATG 258  
 DB 158 TCATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 217  
 QY 259 GCCAGATTTGGCGGCGGACGAGTGGCGCTGCGCGGCGGCGGCGGCGGCGGCG 318  
 DB 218 GCCAGATTTGGCGGCGGACGAGTGGCGCTGCGCGGCGGCGGCGGCGGCGGCG 277  
 QY 319 ACGAGTGGCGCTTCCCGTTCGCGGCGGCTGCTGCAATGAGCGCGGAGGCGCG 378  
 DB 278 ACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 337  
 QY 379 AGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 438  
 DB 338 AAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 397  
 QY 439 GTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498  
 DB 398 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456  
 QY 499 ACTGCGAGTGTGCGGCGGAGTCCATGCTTAC-----GGCGCATGAGCTACGGCGT 551  
 DB 457 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516  
 QY 552 GGAAGTGAACCTTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 605  
 DB 517 GCTTCTGGGCTTACTACCCCATGAGAGTATGCTGCTGCTTAACTCAGAGATCC 576  
 QY 606 CTCCTACCAAGCGGCGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 665  
 DB 577 CTCTTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 636  
 QY 666 TCTTCTATGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 725  
 DB 637 CCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 693  
 QY 726 CCTGTGCAACCCGAGTCTATGAGCCATCCAGAGATCTTCTGATATGAGTATGAT 785  
 DB 694 TCTTGTGCGGCTGAGCCCATGATGATGATGATGATGATGATGATGATGATGAT 753  
 QY 786 GTTGTGGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845  
 DB 754 GTTGTGGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813  
 QY 846 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884  
 DB 814 GTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 873  
 QY 885 CTGCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944

QY 3090 TTATGATACCTCTCAACCACTTCTCTTATGTAACCTCAATGCTGTGCTGCGCTT 3149  
Db 3184 CTCTGATACCCCGACACGCTCTCTCTGTAACCTCAATCGGGGGTGGCGGAGATC 3243  
QY 3150 TCAATTCGATCATATAACGATATGATCATGAGGAGCCCTCTTTGGGAAGTATCTTT 3209  
Db 3244 TCGAACCGCATCAACAGGGGTACGATGCTGGGAGCCCTGTTGGGAAGCTCTTCTTC 3303  
QY 3210 GCATTTGGGTGATTTGTCATCTTTATCCCTTTCTCAAGATTGTTGGAAGCAAAAC 3269  
Db 3304 GCCTTCGGGTGATTCGTCACCTGTACCGTCTCTCAAGGCTGTGGGGGAGCAAGAAC 3363  
QY 3270 AGAACACCAACGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3329  
Db 3364 AGAACCGCGACGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3423  
QY 3330 TGGGTTGGATGATGCTCTCTCTGCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 3389  
Db 3424 TGGGTTGGATGATGCTCTCTCTGCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 3483  
QY 3390 TTGATTCGCACT 3402  
Db 3484 CTGAGCTGCACT 3496

RESULT 4  
AF304374 3847 bp mRNA linear PLN 02-MAY-2001  
LOCUS AF304374  
DEFINITION Nicotiana glauca cellulose synthase catalytic subunit (cesal) mRNA,  
complete cds.  
ACCESSION AF304374  
VERSION AF304374.1 GI:13925880  
KEYWORDS  
SOURCE  
ORGANISM  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 3847)  
Doblin, M.S., De Melis, L., Newbiggin, E., Bacic, A. and Read, S.M.  
Pollen tubes of Nicotiana glauca express two genes from different  
beta-glucan synthase families  
Plant Physiol. 125 (4), 2040-2052 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
11299383  
2 (bases 1 to 3847)  
Doblin, M.S., De Melis, L., Newbiggin, E., Bacic, A. and Read, S.M.  
Direct Submission  
Submitted (12-SEP-2000) Plant Biology, University of California, 1  
Shields Ave, Davis, CA 95616, USA  
FEATURES  
source  
1.3847  
/organism="Nicotiana glauca"  
/mol\_type="mRNA"  
/db\_xref="taxon:4087"  
/rname="p1st11"  
/note="genotype: S282"  
1.3847  
/gene="cesal1"  
103..3378  
/gene="cesal1"  
/product="cellulose synthase catalytic subunit"  
/protein\_id="AAK49454.1"  
/translation="MDTKRLVAGSHNRNEPVVINDVRSVSVKLSQIQICD  
EIEVTVDGEPIFACNECAFVRCQCYERREGQACPOCKTFRKIKGSPVADGDE  
DDEFDDVDFEDYHGNBRYSEAFSRLRGTHNAGLTTSEVDPALNSBEPIL  
TYGQEDDTISADKHALIIPPMGRGKVHPVSDMSLPRLMDPKDLAVYGVTV  
AWKBERMDKKKNDKLOVYKHGGGKNDGDELDPDLKMDREGQPLSRKPLIS  
SRUSPRLILVRLAVGLFPHYRITHPNDAYALMLISICEIWFVSVIPFOPKM  
PIVRETYDLRLRYRKEGKPSGLATIDFVSTVDLPKPLPLTANTVUSILAVDP  
EDKVSCTVSDGAAMLTFEALISTSEFARKWPFCKKFNIEPRAPEWYFSOKVDYLN

gene  
CDS

BASE COUNT 1061 a 719 c 925 g 1142 t  
ORIGIN  
Query Match 38.9%; Score 1531.4; DB 8; Length 3847;  
Best Local Similarity 68.2%; Pred. No. 8-286;  
Matches 2252; Conservative 0; Mismatches 996; Indels 53; Gaps 7;  
DCE"  
KVPSFVREBRAMKRDYEFKRIINGLAVAKYDEPDQTMODTPMPQNLVRDHPGM  
IOVFLNGCVRDI BGNVRLPYVSRERKPGDHHKAGAMALRVASVNAFYL  
INDCOHYINNSALREAMCFMMDPTSGKIK CVQGPQRPDGLDRDRYSNRNVFPI  
NMKGDDIGQPIYVGTGCVFRQALYGDAPAKTGPBGTICMPCMKWCCCGSGSPFV  
KMGKTKDNKPKTKTEASPOLHALENIIEGIDSEKATLMPDIXLEKKGQSPVF  
VASTLEDDGIPGATSAIILKEAHLVISCQEDTEMRGVMTYGSVETDILNGPK  
MACHMRSVYCMKPRKAPKSGAPILUSBRHOUVRMALSVETILSKSPRIYVGGCC  
LKPFRFSTINSVTPPLSLPLIATCAVAVLLTGKTFVPEISNYSVLENGFLIMI  
AATSVLEMQMGVTTIDDMWRNEQFVYIGGASHPLAFQGLKVLAVGSTPFTVSKA  
ADDFSELSYLFKMTSLIPWTLIINIIVIVGIVISDPAINGGSDSWGFLPFLAL  
NVIIVHLYPLKVMGRONKVPITIVVMSILASIFSLMLVRVNPFRAGVLVLEVCGL  
DCE"  
QY 139 AAGCATGAGAGGCGAGCGCGGCTGTGTGCGCGCTCCACAAACGCAAGCTGTG 198  
Db 98 AAGTAATGATCTAATAAGGAGGCTAGTTGCTGTTCACACATAGAAATGATTTGTG 157  
QY 199 TCATCCGCGCGAGCGCGATCCCGCGCGAGCGCGGAGAGCAAGACGGCGAGCT 258  
Db 158 TCATCAATGCTGATGATGTTGAGAGATTAATCTTGTGAAAGATTAAGTGGCGAGATT 217  
QY 259 GCCAGATTTGCGCGAGCGAGCTGCGCTTGCCTCCCGCGCGGAGACCTTGTGTGCGTGA 318  
Db 218 GCCAGATTTGTGAGAGATGAGATTGAAGTACAGTAGATGAGGACCATTTATGCTTGA 277  
QY 319 ACGAGTGGCTTCCCGCTGTGCGCGGAGCTGCTAGCAATACAGCGCGGAGGACCGC 378  
Db 278 ACGAATGATCATCTCCCTGTTGCGAGCAATGCTATAGATTAAGAGAAAGAAATTC 337  
QY 379 AGAATGCGCGCGAGCTGCAAGCTGATACAGCGCTCAAGGCTGCCAAGCTGACCG 438  
Db 338 AAGCTTGCCTCAATGCAAACTCGGTCAAGGCAATTAAGGAGACCCGAGGTTGATG 397  
QY 439 GTGACGAGAGAGAGAGCGCGCTGATGACCTTGACACAGATTCACCTGGAACGCGCATG 498  
Db 398 GAGATATGAGATGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 456  
QY 499 ACTCGAGCTGTGCGCGAGCTGCTTAC-----GCCCATGAGCTACGCGCGT 551  
Db 457 CCTCGATATATGCTGAGGCTGATCTCTCTGCTGCTGCGCGCGGATCAATCATTAAT 516  
QY 552 GGAAGTGACCTTAATGCGCGGACAG-----CTTCCAGTCAACCCCAATGTTCCA 605  
Db 517 GCTTCTGGGCTTACACCCCATCAAGATGATCTGCTTAACTCAGAGATCCCT 576  
QY 606 CTCCTACCAACGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 665  
Db 577 CTCTTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 636  
QY 666 TCTTTCATGAGGTGTGGGAGAGAGATCATCTCTTCTTATGCGATCCAGCTTA 725  
Db 637 CCACCGTTATGAGGTGCGGAGAGAGAGATGATCTGCTGCTTATCTGATTCGA---TG 693  
QY 726 CCGTGCAACCCAGGTCTATGACCCCATCAAGGATCTTGTGCAATATGATGATGATG 785  
Db 694 TCTTTCCTCCCTGACCCATGATCTTAAGAAAGATTTGGCAATTAATGATGATGATG 753  
QY 786 GTTCTTGAAGAAACGATGGAATTTGAAAGCAAGAGAGAGAGATGACAGACG 845  
Db 754 GTTGCATGAGAGAGAGAGATGAGAGCTGAGAGAGAGAGAGAGAGATGATGATGATG 813  
QY 846 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884  
Db 814 GTTAAACGAGAGAGCGGTAAAGGTGTGTAACGATGAGATGATGATGATGATGATGAT 873  
QY 885 CTACCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944

Db	874	TTGCCAAGATGATGAAGGTAGACAAACCACTTCGAGAAAATGCTCATTTTCTCTACG	933
Qy	945	CAGATTAATCCATATAGATGATATATATATTCGCGCTTGATGTTGGGGGTTCTCTTC	1004
Db	934	AGCGTAAGCCCATACAGATGTGCTATTTTAGTTCGTCCTTGCTGTGGGCTCTTTTTC	993
Qy	1005	CACTAACGATGATGCATCCGGTGAATGATGCAATTTGCTTGCGCTCATATCTGTATC	1064
Db	994	CACATATAGAAATACCATCTGTCTCATATGATGATATGCGTGTGGCTATTTCTATATA	1053
Qy	1065	TGTGAATCTGGTTTGGCCATGTCCTTGATCTTGTATCAATTCGCCAAGTGTCCCTATT	1122
Db	1054	TGTGAATATGTGTGGGTGCGTATCAAGATATTTGATCAAGTTCACAAAATGGTTCCGATT	1113
Qy	1125	GAGAGAGACCTTACCTAGACCGCGCTGTCACCTGAGGTTCCGACAAAGAAAGCCAGCCACT	1184
Db	1114	GTGCGAGAGACATTAAGCTGACAGGCTATCTCTGAGGTATGAGAAAGAAAGGACCTTCT	1173
Qy	1185	CAACTGCTCCAAATTAATTTCTTTGTCAATACGTTGATCCCTTAAAGAACTTCCTTGTG	1244
Db	1174	GGGTTAGCCCCCTATAGATATATTTGTATGACGGTGGATCCCTGAAAGAACTTCGCTT	1233
Qy	1245	GTCAACAACAATACCTGTTCTATCTATCTCTTCCGTGATATCTCTGTTGATTAAGTTCCT	1304
Db	1234	ATTACTGCAAAATACGTTCCTCCACTTCCTGCTGATATATCCGAGAGACAAAGTTCCT	1293
Qy	1305	TGCTATGTTTCTGATGATGATGTCGTGCATATGCTATGCAAGTTTGAAGCATATCTGAACATCT	1364
Db	1294	TGCTATGTCCTGAGAGATGGTGTCTGCATGCTTACTTTGAGGCTCTCTGAAACATCT	1355
Qy	1365	GAATTTGCAAGAAAATGGGTTCTCTTCTGCAACCGTACAAATATTTGAACCTTCGCGTCCA	1424
Db	1354	GAGTTGCGAAGAAATGGGTCCACTTCGCAAGAAAGTTCAACATATAGACCTTCGGCCCCA	1413
Qy	1425	GAGTGTACTTCTCAACAGAGATAGACTTCTTGAAGAACAGAGTGGCACCAACTTGTGT	1484
Db	1414	GAGTGTATTTCTCTCAGAGGTTGATTAATCTGAAAAATTAAGTACATCCATCAATTTGTG	1473
Qy	1485	AGCGAGAGAGCAATGAAAGAGATGATGGAATTCAGAGTCAGATCAATCAAGCTCTTA	1544
Db	1474	AGGGAACGTCTGTCTATGAGAGACATATGAAATTCAGAGTTCCGATTAATATGACTG	1533
Qy	1545	GTTCGCAAAAGCCACAGAAAGTCTCTGAGAAGAGATGACAAATGCAAGATGAAACCCCTGG	1604
Db	1534	GTTCGCAAGGACACAAAGGTTCTGAAATGGCTGACCAATGCAAGATGGAACCGCTTGG	1593
Qy	1605	CCTGGAACAATGTTCGTGATCATCTCTGGAATGATTCAGGTCCTCTTGGCCAAAGCGGA	1664
Db	1594	CCCGGAAATCTTGTACAGGATCATCTGGAATGATCCAGGTCCTCTGGGTAAATGATGAT	1653
Qy	1665	GCGCTTGACTGTGAGGGAATGCAACTGCCACGATTTGTTATGTTTCTTAGAGAGAAAGGA	1724
Db	1654	GTCCGTGATATTTGAAGGGAATGTACTGCTCCCTTAATCTATGTTTCTCGAGAAAGGT	1713
Qy	1725	CCAGGCTATTAACCATCATATAGAAAGCTGGTGTATGAATGCAATGTCGAGTCTCTGCT	1784
Db	1714	CCAGGATTTGACACACCAAGAAAGCTGTGTGCATGAATGCTTTGATGCGGGATACAGCG	1773
Qy	1785	GTACTTAACAAATGCTCCATATTTGTTAAACTTGATGTGATCACTACATCAACAACAGC	1844
Db	1774	GTCAATATCAAAATGCTCTTACTTGTCTCAATGTGATGTGATCACTATATATAATTAACGT	1833
Qy	1845	AAGGCTATTAAGGAAGCAATGTGTTTATGATGACCTTTAATAGGAAAGAAAGTGTTC	1904
Db	1834	AAGGACTGAGAGAAAGCTATGTGCTTTATGATGAGCCCACTTCAGGAAAGAAATATGTC	1893
Qy	1905	TATGTACAGTTTCCCTCAAAAGATTTGATGGGATTTGATCCGCATAGACGATATGCTAAACGG	1964
Db	1894	TATGTACATTTCCCAAAAGGTTTATGTGATTTGACAGCATGACAGATATCTCAACCGC	1953
Qy	1965	AATGTTGCTTTTGTGATATCAACAATGAAGGTTTGATGTATTCAGGGTCCAAATTTAT	2024
Db	1954	AACGTGCTTCTTTGATATCAATATGAAGGACTTGAACGGGATCCAGGGTCCAAATTTAC	2013

QY	2025	GTGGTACTGATGTGTAATTTAAGAGCAGGCAATTATATGGTATATGATGCCCCAAACA	2084
DB	2014	GTGGGTACTCGATGTGTCTTCAGAGAGGACAGCCGCTTATATGGTATATGATGCTCCAAAGAG	2073
QY	2085	AAGAAGCCACATCAGAGACTTTGCACTGTGGCCCAAGTGTGCTTTTGTGTGTCTGC	2144
DB	2074	ACAAAACCTCCAGGAAAAACATGCACTGTGGCCCAAAATGTGTGTGCTTGTGCTGCT	2133
QY	2145	TTTGGCAATAGGAAGCAAAAGAACTACAAACCCAAAACAGAGAAAGTAATATTA	2204
DB	2134	TCT-----CGAAAGAACATTAAGAAAGGAAACACAGAGATACAAAG	2181
QY	2205	TTTTCAAGAAAGAGAACCAATCCCTGCAATATGCTCTTGTGTAAATTAGCAAGCT	2264
DB	2182	AAGACTAAACCAAGAGGCTTCCACCAAAATTCAGGCCCTTGAATAATATCGAAGAAAG	2241
QY	2285	GCTCCAGAGAGCTGAGAAATGAAAAGCCGGATTTGTAAATCACAAAAATTGAAAAGAA	2324
DB	2242	ATTGAGGAATGTATAGTAAAAAGCAACCTCATGCTCCAGATMAAATTGAAAGAA	2301
QY	2325	TTTGGCCAACTTCGTGTTTTTTGTAACATCCACTTCGGAATGGAGAACCTTGAAG	2384
DB	2302	TTTGGACATATACAGATTTGTGTTCTTCAACACTTTCAGAAATGGGCAATTCCTCA	2361
QY	2385	AGTCAGACTCCGTGCTTCTTTTGAAGAAGCTATACATGTCATTAGTGTGGTTAGAA	2444
DB	2362	GGAGCAATCATGAGATCTCTCTGAAAGAACATACATGCTATTATGTGTGTTATGAA	2421
QY	2445	GACAAAGACAGCTGGGGAAAAAGATTGGCTGATCTATGGATCAGTTACAGAAATAT	2504
DB	2422	GACAAACAGATGGGGTAGAGAGGTTGGAATGATTTATGGATCTGTTACTGAAGAAATC	2481
QY	2505	CTAATGGTTCAAGATGCAATGTCATAGTGGGGGCAATTTACTGCATACCTAAACG	2564
DB	2482	CTAACGGATTTAAGATGCACTGTGACGGTTGGAGATCTGTACTGATGCCCAAGG	2541
QY	2565	GTTGCATTCAAGGTTGTGCACCTGTGATCTTTCAGATCGTCTTCAACAGAGCTTCGG	2624
DB	2542	CCTGGTTTAAAGGGTCAGCTCTCATATTTCAATCTTCAATCGTCTTCAACAGGTTCTGCG	2601
QY	2625	TGGGCTTGGGCTCATATTGAGATCTTTCAGCAATCATGCGCTCTTGGTATGGGAT	2684
DB	2602	TGGGCTTGGGATCTGTGAATTCCTGTGAGTAAACATTGTCCCATTTGGTATGATAT	2661
QY	2685	GGTGGCGGCTCGAAATTTTGGAAAGATTTTCTCAACATCAACTCATTACGTATCTTGG	2744
DB	2662	GGGTGTGGGCTGAAGCCGCTGGACGCTTCTTCAATAAACCTAGTGTCTACCCATTA	2721
QY	2745	ACATCTATATCCCTCTTGGCTTACTGTACATATGCTGTCGATCTGTTTATTGACAGGGAA	2804
DB	2722	ACTTCTCTTCTTGTGATGCAATATTTGTCATTTGTCAGAGCTGTCTTGTCTTACGGGAA	2781
QY	2805	TTTATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTCATGTCACTTTTATCTGC	2864
DB	2782	TTTATATGTTCCAGAGATTAGCACTAATGTCTATGATCTTTTCAAGGGACTTTTATATG	2841
QY	2865	ATTTTGTCTACAGACATCTTGAATAAGATGAAGTGTGTGGAATTTGATGACTGTGG	2924
DB	2842	ATTGCTGTCAAGAGTGTCTTGAATGCAATGGGAGGTGTTATCTATTTGATGACTGTGG	2901
QY	2925	AGGAATGAGAGTTCTGGGTGATTTGGAGGTGTGTCTCAACCTCTTGTGTGTCCAG	2984
DB	2902	AGAAACGAGAGTTCTGGGTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2961
QY	2985	GGACTTTCAGAGCTATAGCTGTGTGTGATCAAGCTTCAACCTGATCAAAAGGTGGA	3044
DB	2962	GGTCTCTTAAAGTTTGGGCTGGGTGACGACTATAGCTTCACTGTATCTTCCAAAGCAGA	3021
QY	3045	GATGATGAGAGTTCTCAGAGCTATATATCATTTCAATGAGACTATATTGATATCTTCT	3104
DB	3022	GACACAGGGAGTTTCTGAGCTTATCTGTCAAGTGAGCACTCTTGTGTGATTCCTCT	3081









Arabidopsis (Accession No. AF062485) (PGR98-114)  
 Plant Physiol. 117 (3), 1125 (1998)  
 2 (bases 1 to 3444)  
 JOURNAL  
 TITLE  
 AUTHORS  
 Direct Submission  
 Submitted (01-MAY-1998) Plant Biotechnology Research Center, School  
 of Forestry and Wood Products, Michigan Technological University,  
 Houghton, MI 49931, USA  
 Location/Qualifiers  
 1.3444  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /clone\_lib="ABRC Number CD4-16"  
 <1.3247  
 /codon\_start=2  
 /product="cellulose synthase"  
 /protein\_id="AAC29067.1"  
 /db\_xref="GI:3135611"  
 /translation="RRRLIAGSHNRREFLVIMADENARIRSVQELSGQTCQICRDEIE  
 LTVDEPVCNCAFPVCRPCYERRENOACPOCTKFKELKGSPRVEGEEDD  
 IDLDNFEYGNNGIDVOYSEGISRRSGSPQSLDAPGSGQPLTYDEDE  
 ISSDRALIVPELGHGHRVHPVSLSDPTVAHRRLMVPQKDLAVYGSVAMKDM  
 BEKKRKNELQVYRHGDPDFEDGDDADFPMDDEGRQPLSMKIPKSKINRYRLI  
 VLRLVILGLFPHYRIHPVDATLMLISVCEIWEPAVSLDQFPRMPTRETYLD  
 RLRLREKEKPSGLSPVDVFSITVDPLKPPILATVTSILAVDPVAKVAYSD  
 DGAAMLFELALSETAFARKVPCCKYCEPRAPEYFCHKDYLKMKVPAFVER  
 RAMKRDYEEFKVINALVATOKVPEDEGMODGTPMPKGLCPGSRKRRKATYAADK  
 RDVENNELPRLVVSREKRPQPHKKKAGANSLIRVSGVLNAPVILANDCDHYIN  
 SKALREACMPMDPOSGKICYQFPORPGIDRHRYSRNRYNPPIDIMKGLDGLG  
 PIYGTGCVFRQALYGFDAPKKKGRKTCNCPKCLCFSGSRKRRKATYAADK  
 KOREASQKIHLENIIEBGRKLVNOSTEAMQMLQKRYQSPVPHLQVPLSGDV  
 MANNASPOCKLEKAIQVISRVEDKTEMGVEIGMIVSVEIDILTGKMSHMRHY  
 CTPLAFKGSAPINLSDRLHVOYLMALGVEIFLSHCPIWYGGGLKMLRLSYI  
 NSVYPTSLPIVYCSLPAICLTGKEIYPRISNYASILFMLFSLIATGLIEMOM  
 GRKVIDMWRNEQFVWIGVSAHLFALFQGLKYLAVDPTNTVTSKADDSFDLY  
 LFKWTSLLIPMTLLIINIVGIVGSDAISNGDSGRLFGRLFFMLVITHLRPL  
 KGLKQDMRMPITIVVSLIASILTLWVRNVPFAKGGPILIEICGLDCL"

BASE COUNT 903 a 664 c 860 g 1017 t

ORIGIN

Query Match 37.9%; Score 1489.8; DB 8; Length 3444;  
 Best Local Similarity 67.3%; Pred. No. 9.8e-278;  
 Matches 2189; Conservative 0; Mismatches 1032; Indels 30; Gaps 5;

QY 160 GGCCTGGGCGCGGCTCCCAACGCAACGAGCTCGTCATCCGCGGAGCGGCGATC 219  
 DB 9 GGTTAATGCGCGGCTTCACACAGAAATGATTTGCTCTATTAATGCCGATGAAATG 68  
 QY 220 CCGGCGCGAAGCCGCGCGGAGAGACGAGCGAGTGTGCCAGATTTCGCGCGACGACG 279  
 DB 69 CCGGAATTAAGATCACTGCAAGAGCTGAGTGAACAGATGCAAACTGCGAGAAATAGA 128  
 QY 280 TCGGCGCTTGGCCCCCGGCGGAGACCCCTTCGTGCGTCGACAGAGTGGCTTCCCCGTCT 339  
 DB 129 TCGAATGACTGTTGATGAGAAACGTTTGTGGCATGTAACGAATGTGCATTCCTCTGT 188  
 QY 340 GCGCGGAGCTGCTAGAAATGAGAGCGCGGAGGAGCGAGCAACTGCCCCCAGTGAAGA 399  
 DB 189 GTAGACCTTGTCTTAGAGTACGAAAGACGAGAGGCAATCAAGCTTGTCCACAGTGCAAA 248  
 QY 400 CTGATACAGAGCGGCTCAAGGCTGCAACGCTGTGACCGGTGACGAGAGAGAGAGACGCGG 459  
 DB 249 CCGGTTTCAAGCGTCTTAAGAGAGTCCCAAGAGTTGAAGGTGATGAAGAGAGATGACA 308  
 QY 460 TCGATGACTGAGCAAGAGTTCAACTGGGACGCGCATGACTCGCACTGTGTGCGCGAGT 519  
 DB 309 TTGATGATTAGCAAGAGTTTGA--GTATGGAATAATGAGATTGAGATTATGACG 365  
 QY 520 CCATGCTCTACGCGCATGAGTACGCGCGTGAAGTGAACCTAATAGCGCGCACAG 579  
 DB 366 TTTCTGAAGGATATGCTCTCTCTGTCGCAACTCGGCTTTCCCAACATCTGATTGATT 425

QY 580 CTTTCAGCTCAACCCCAATGTTCCACTCTCTACCAACGCGGAAATGTGATGATCTCC 639  
 DB 426 CAGCTTCACCTGAGCTCTCAGATTCCATTGCTTACCTTACGCGGACGAGAGCTTTGAGATT 485  
 QY 640 CACCGGAGAGCAGCGCG---TGTCGCTCTTTTCACTGGGTGTGGGAGGAAAGAGATAC 696  
 DB 486 CTTTGTATGACATGCTCTTATTTGTTCTCTCTTCACTTGTGTGTATGCAATATAGAGTTC 545  
 QY 697 ATCCCTCTCTTATGCGGATCCAGGTTACTCTGTGCAACCCAGTCTATGACCCATCCA 756  
 DB 546 ATCTGTTTCTCTTCTGACCCGACCGGTGTCGACATGAAAGGCTGATGTATCTCACA 605  
 QY 757 AGGATTTTCTGATATGGGTATGTGATGTGTTGTTGAAAGGACGATGAGATTTGCA 816  
 DB 606 AAGATCTTCGGTTATATGTTATGAAAGTGTGCTTGAAGAAATCGATGAGAGATGGA 665  
 QY 817 AGCAGACAGAGAGAGATGACACGAGCGGCA-----ATGATGTTGGTGGTATG 867  
 DB 666 AGGAAAGCAAGATGAGAACTTCAGGTTGTTAGGATGAAGAGATCTGATTTTGAAG 725  
 QY 868 ATGCTGACGATGCTGATCTTACCACTAATGATGAAGCAAGCAACTGTCGAGAAAA 927  
 DB 726 ATGCTGATGATGCTGATTTTCCAAATGATGATGAGGAAAGGACCATTTGTCTATGAAG 785  
 QY 928 TTCCACTTCTCAACGCAAGTTAATCCATATAGATATATATATATATATATATATATAT 987  
 DB 786 TACCAATCAAACTGACCAAGATTAATCTTACCGATGTTAATGTGCTACGCTTGTGA 845  
 QY 988 TTTTGGGAGTCTTCTTCCACTACCGAGTATGATGATCGGATGATGATGATGATGATGAT 1047  
 DB 846 TTTTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 905  
 QY 1048 GGCCTATATCTGTTATCTGTAATCTGTTGATGATGATGATGATGATGATGATGATGAT 1107  
 DB 906 GGTCTATTTCTGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965  
 QY 1108 CAAGGTGTTCCCTATATGAGAGAGACCTTACTAACCAGGCTGCTACGAGTGTGACA 1167  
 DB 966 CTAAATGTATACCTTATGACGAGCAAGCTTATGACGAGCTCTATTAAGATATGAGA 1025  
 QY 1168 AGGAGGCGCAGCATCTCAACTGCTCAATGATTTCTTGTCACTAGCGGTGATCCCT 1227  
 DB 1026 AAGAAAGGAAACCTGTGGGACTATCCCTGTGATGATTTGTTAGTACAGTGGATTCAT 1085  
 QY 1228 TAAAGAACCTCTTGTGTACACAACAATACTGTTCTATCTATCTTGTGGTGAATATC 1287  
 DB 1086 TGAAGAGCCTCGCTTATTAACGCAAAATAGTGTCTGCTATCTTGTGATGATATC 1145  
 QY 1288 CTGTTGTAAGTTTCTGCTATGTTCTGATGATGATGATGATGATGATGATGATGATGAT 1347  
 DB 1146 CTGTGATTAAGGTGTGTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1205  
 QY 1348 CATTATCTGAAACATGAAATTTGCAAGAAATGAGTTCCTTCTGCAAAACGCTACAA 1407  
 DB 1206 CTCTTTCTGAGACCGCTGAAATGCAAGAAATGAGTTCCTTCTGCAAAATATTTGA 1265  
 QY 1408 TTGAACCTCGCGCTCCAGAGTGTACTTCCACAGAAATGACTACTTTGAAAGACAGAG 1467  
 DB 1266 TTGAGCCTGTGCTCCGAAATGATTTTGTGCAATTAATGACTACTTTGAAAGATTAAG 1325  
 QY 1468 TGGCAGCAACTTTGTTAGGAGAGAGAGCAATGAAGAGATATGAGAAATTCAGAG 1527  
 DB 1326 TTATCCCGGATTTGTTAGGAGAGCGGAGCCATGAAGAGATTTTGAAGATTCAGAG 1385  
 QY 1528 TGAGAAATCAATGCTTAGTTGCAAGAGCCAGAAAGTCTGAAAGAGATGACATG 1587  
 DB 1386 TAAAGATCAATGCTTTAGTACGAAACAGACAGAAATGCTGTGAGATGATGATGATGATG 1445  
 QY 1588 AAGATGAAACCCCTGCGCTGAAACATATGTTGTATCAATCTGTAATGATTCAGGTCT 1647  
 DB 1446 AAGCGGTACACCTTGCGCGGTAATATGTTGTGAGATATCTTGCAATGATTCAGGTCT 1505

1648 TCCTGGCCAAAGCGAGGCTTGAAGTGGAGGAAATGAATGACGACGATGGTTATG 1707  
 1506 TCCTGGAGTACCGGTGCTGCTGATGTGAAAAACAGAGTTGCTGATTAAGTTACG 1565  
 1708 TTTCTAGAGAGAAAGCAGCGCTATTAACCATATTAAGAAAGCTGGTGTATGAATGCAT 1767  
 1566 TTTCTCGTAGAAGAGACCGGATTTGATCACCATTAAGAAAGCTGGAGCTATGAATCCC 1625  
 1768 TGGTCGAGTCTCTGCTGATTAACAATGCTCCATATTGTTAACTTGAATGTGATC 1827  
 1626 TGATCGAGTCTCTGGGCTTATCAAAATGCTCTTACCTTGAATGCGATTGATC 1685  
 1828 ACTAATCAACAACGAGCGCTATTAAGAAAGCAATGTTTATGATGAGACCCCTTAC 1887  
 1686 ACTAATCAACAATGAGCAAGCTCTTAAGAAAGCAATGTTTATGATGATGATCCTCAGT 1745  
 1746 CAGAAAGAAATCTGTTATGTTCAAGTTCCCTCAAGGTTGATGGATTTGATAGGCAAG 1805  
 1948 ACCGATATGCTAACCGGAATGTTGCTTTTGTATTAACAATGAAGGTTGATGGTA 2007  
 1806 ATCGATACTCAATGCAATGTTGCTTGTGATTAATGAAGGTTGATGGAGGC 1865  
 2008 TTCAGGTCCTAATTAATGTTGATGATGATGATTAAGAAAGCGGATTAATGAT 2067  
 1866 TACAAGGCGCTATATACGTGGTACAGTTGTTTTCAGAGGCAAGGCTTTACGAT 1925  
 2068 ATGATGCCCCCAAAACAAAGAAAGCCATCAAGAGCTTGAACCTGGCCCAAGTGT 2127  
 1926 TTGATGACCGAAGAAAGAAAGGCGCCATGAAGCATGCAATGCTGGCCAAATGAT 1985  
 2128 GCTTTGCTGTTGCTGCTTTGGCATAGAGCAAAAGAAAGCTACCAAAACCAAG 2187  
 1986 GTCTCCTATGTTGGTTCAAGAAAGATGTTAAACCAAGAGAGGCTGGCGATGA 2045  
 2188 AAAAAAGAAAGTATTTATTTTCAAGAAAGAAAGCAATCCCTGATATGCTTG 2247  
 2046 AAAAAAGAA-----TAAAGAGCGCTCAAGAGATCCAGCATTAAGAAATA 2093  
 2248 GTGAATTAAGAAAGCTGCTCCAGAGCTGAGATGAAGAAAGCGGATTTGTAATCAAC 2307  
 2094 TCGAAGAGGCGCGGCTCATAAAGTTCTTAAGTAAAGCAAGTCAACGAGGATCCAAA 2153  
 2308 AAAAAATTAGAAAGAAATTTGGCAATCTTCTGTTTGTATCATCCACATTTCTGAGA 2367  
 2154 TGAAGTTGCAAGAAATATGGGCACTCTCTGATTTGTTGATCTGGCGCTGAGAA 2213  
 2368 ATGTTGGAACCTTGAAGAGTCAAGTCTGCTCTCTTTGAAAGAAAGCTATACATGCA 2427  
 2214 ATGTTGGGATGCTTAAGAAAGCAAGCGCGCTTGTCTGCTTAAAGAAAGCATCCAGTCA 2273  
 2428 TTAGTTGTTATGAAGCAAGACAGACTGGGGAAGAAAGATGGCTGATTAAGAT 2487  
 2274 TTAGTCCGATATGAAGATTAAGTGAATGAGGAAAGAAAGATTTGGTGAATTAAGT 2333  
 2488 CAGTTACAGAAATATTTCAATGCTTTCAAGATGATGTCATGTTGGCGTCAATT 2547  
 2334 CTGTTACGAGATATTTCTTAAGGTTTCAAGATGATTTCTATGTTGAGACATGTTT 2393  
 2548 ACTGATTAACCTTAACGGGTTGCTTAAAGGTTCTGACCTCTGAATCTTTCAATGCTC 2607  
 2394 ATTTGACACCAAGTTAGCGGCTTTCAAGAGATCACTCCAAATCAATCTTTGCGATGCTC 2453  
 2608 TTCACAGGTTGCTGGTGGGCTCTGGGCTATTAAGATCTTCTCAGAAATCATTTGCG 2667  
 2454 TCCATCAAGTTCTTCATGAGGCGCTTGGGCTGGGTTAGATTTTCTGAGTAGGCAATGCTC 2513  
 2668 CTCTTTGATGAGGATGAGTGGCGCTGCAATTTTGGAAAGATTTTCTTACATCACT 2727  
 2514 CTATTTGATGATGATGAGGAGTGAAGATGCTTGAAGCGGTTGCTTACATTAAC 2573  
 2728 CCATGCTGATCTTGAACATCTATTCCCTCTTGCTTACGTAATGCTGCACTCT 2787

2574 CTTGATTTTACCCGAGAACCTCTACCGCTCATGTTTATGTTCTCTCCCTGCACT 2633  
 2788 GTTTATTGACAGGAAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTTCA 2847  
 2634 GTCTTCTACATGGAATAATTCATCTGTTCCGAGATGAGAACTATCCGATCTCTTCA 2693  
 2848 TGTCACTTTTATCTGCAATTTTGTAGAGACCTTAAGAAATGAGATGAGTGTGTTG 2907  
 2694 TGGCGCTCTCTCGATGATGCAATACGAGGATTTCTGAGATGCAATAGGGAAGTTG 2753  
 2908 GAATTAAGTACGTGTGAGAAATGAGCAGTTCTGGGCTCATTTGAGAGTGTCTCACAC 2967  
 2754 GATGATGATGTTGGTGAAGAAACGACGTTTGGTGTATGAGAGTGTCTTGCGCATC 2813  
 2968 TCTTTGCTGTGTTCCAGGACTTCTCAAGTTCATAGCTGTGTGATTAACAAGCTTACCG 3027  
 2814 TGTGCTCTCTTCCAGAGGCTCCCAAGGTTCTTGATGATGATGACATTAACCTTACAG 2873  
 3028 TGACATCAAGGCGGAGATGATGAGAGGTTCTGAGCTATATACATTAATGAGCTA 3087  
 2874 TCACATCAAAAGACAGCTGATGAGAGGTTCTGAGCTTATCTTCAATGAGCTT 2933  
 3088 CCTTATTAATCTCTCTCAACCTTGTCTTATTAATGAATTCATTTGATGATGATGAG 3147  
 2934 CACTTCTCATCCCTCAATGACTATCATATCAATTAAGTCAATGAGTCAATGCTGAG 2993  
 3148 TTTCAATGCGATCAATTAACGATATGAGTCAATGAGGCGCCCTTTGGGAAGCTATTCT 3207  
 2994 TCTTGATGCGATCAGCAATGAGATGAGACTGATGAGGAGCCGCTTTTCGAAGACGTTCT 3053  
 3208 TTGCAATTTGGGATATGTTCCATCTTTATCCCTTTCTCAAGGTTGTTGGAAGCAAA 3267  
 3054 TTGCACTTTGGGATCAATTAATCTTACCCGTTCTTAAAGGTTGCTTGGAAACAAG 3113  
 3268 ACAGACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3327  
 3114 ATGAATGCAACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 3173  
 3328 TTTGGGTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3387  
 3174 TTTGGGTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3230  
 3388 GTTTGATGCTC 3398  
 3231 GTTTAGACTGC 3241

RESULT 7  
 AF027173  
 LOCUS 3771 bp mRNA linear PLN 07-FEB-1998  
 DEFINITION Arabidopsis thaliana cellulose synthase catalytic subunit (Ach-A)  
 AF027173  
 VERSION complete cds.  
 KEYWORDS AF027173.1 GI:2827140  
 SOURCE  
 ORGANISM Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE  
 1 (bases 1 to 3771)  
 Artoli, T., Peng, L., Betzner, A. S., Burn, J., Witke, W., Herth, W.,  
 Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J.,  
 Redmond, J. and Williamson, R.B.  
 Molecular analysis of cellulose biosynthesis in Arabidopsis  
 Science 279 (5351), 717-720 (1998)  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 9445479  
 2 (bases 1 to 3771)  
 REFERENCE  
 Artoli, T.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (29-SEP-1997) Plant Science Centre, Australian National

FEATURES  
source University, Acton, Canberra, ACT 200, Australia  
Location/Qualifiers  
1..3771  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
1..3771  
/gene="Alc-A"  
/gene="Alc-A"  
217..3471  
/note="RSM1-like"  
/codon\_start=1  
/product="cellulose synthase catalytic subunit"  
/protein\_id="AAC39335.1"  
/db\_xref="GI:2827141"  
/translation="MNTGRLIAGSHNRNFVLIINDESARISRVQELSGTCOICGD  
EIELTSSSELFVACNECAPVPCRCYEVREBREGNOACPOCKRYRIRKISPRVGDDE  
BEEDIDLEVERPDHGMDEPHAAALSSRLNTGRGLDAPRSQIPLITTCDEADAM  
YSDRHALIVPSTGKNRYPAFTDSAPPOARSVPQIDLAIEYGSVAMDKREIV  
WKRROEKLQVIRHGGNNRGSNDDELDDPMMDGROPLSRKPIRSRIPIY  
RMILCRALILGLFPHYRLHPVNDYGLMTLSVCEIWFVWSIMIDQPKMYIRE  
TYDLRLSLRYEKSGPSGLAPVDFSTVDPLKEPILANTVLSILADVDPDKAC  
YVDDGAAALTEPRLSDTAEPARKWPKCKNIIPRAEPWFSOKMDLKKVHPAF  
VRRRAMKRDYEEFKIKINALIYATKAVPEEGTMDGTPEPGRNTRDHPM1QVPLG  
HSGVRIDGNEELRYVSRERKPGDFHKKAGAMNSLIRVSALVSNAPYLINVDCH  
YINNSKAIRESMCFMMDPOSKKVCYVOPORPDGIDRDRSNRNVFEDINMGDL  
GIQGIYVGCVFRRQALYGFDAPKKKRPGKTCNMPKMCCLCGLKKSKTKAKD  
KKTNETSKOIHLENDVEGIVPVSNVSRSEATOLLEKFGSPVFAVAVON  
GVPRNASPACILIREALIOVSCGVEDTKMGKIGWIVGSVEDILTGKMGHGRS  
VYCMPRKAPFKSAPINLSDRLHQLVRLMAGSVETLSHRCPIWYGGGLKRLERS  
YINSVVPMWTSPLIYCSLPVCLLTGLFVEIISVNGIIFMLPFI1AVTGLEM  
OMGVGIDDMWRNEQFVIGASSHLFAFOGLKVLAVNTMFTVTSAAADPASE  
LYFKWTLILPPTLLINIIGIVGDAISNGDSMGPLFGRLEFALWIVLHYLP  
FLKMGKDKMPTIIVVMSILASITLLMWVNFVAKGPFVEICGLNGCN"

BASE COUNT 1037 a 699 c 907 g 1128 t  
ORIGIN

Query Match 36.4%; Score 1434.2; DB 8; Length 3771;  
Best Local Similarity 65.9%; Pred. No. 5,7e-267;  
Matches 2175; Conservative 0; Mismatches 1063; Indels 63; Gaps 4;

QY 124 GAAGTGGAGGGGGAAGCATGAGAGCCGCGGCTGTGGCCGCTCCCAACACC 183  
DB 197 GATTGCGGTGAGGAAGCATCATGAATACTGGTGTGGCTCATTTGCTTCACAAACA 256  
QY 184 GCAACGAGCTCGTGCATCCGCGAGAGCGCATCCCGGCGGAAGCCGCGGGAGC 243  
DB 257 GAAACGAAATCGTTCTCATTAACGCCGATGAGTGCAGAAATACGATCACTGTAAGAAC 316  
QY 244 AGAACGGGAGGTGTGCAGATTTGCGGCGAGCAGCGTGGCTTGGCCCCCGGGGAGCC 303  
DB 317 TGAAGTGGGCAAAATGTCGTAATCTGTGGAGATGAATCGAATTAACGGTTAGAGAGAGC 376  
QY 304 CTTTGTGGCGTCAACGAGTGGCTTCCCGCTGTGCGGAGCTGTAAGATACGAGC 363  
DB 377 TCTTTGTGCTTCAACGAAATGCGCAATCCGGTTTGTAGAACCATGATGATGATGAAAC 426  
QY 364 GCGGGGAGGCGAAGAACTGCCCGCAGTGAAGACTGTATCAAGCCGCTCAAGGGCT 423  
DB 437 GTAGAGAGAAATCAAGCTTGTCTAGTGCAAAACCTGATCAAAAAGGATTTAAAGGTA 496  
QY 424 GCCAACGTGTGACCGGTGACGAGAGAGAGACGGCGTGCATGACCTTGACAAAGATTTCA 483  
DB 497 GTCCACGGGTTGATGGAGATGATGAAGAAAGACATTTGATGATCTTGAATGAGT 556  
QY 484 ACTGGAGAGCGCATGACTGCCAGTCTGTGGCCGAGTGCATGCTTACAGGCAATGAGCT 543  
DB 557 TTGATCATGGAGTGAACCTTGAAACATGCGCTGAAGCGCACCTCTTCAACGCTTTAAACA 616  
QY 544 ACCGCGGTGAGGTGACCTTAATGAGCGGCGCAAGCTTTCCAGCTCAACCCCAATGTTG 603  
DB 617 CCGGTGTGTGATTTGATTTCAAGCTTCACCTGGCTCT-----CAAGTTTC 661

QY 604 CACTCTCTACCAACGGGCAAAATGTTGATGATGATCCACCGGAGACGACCGCTGTGC 663  
DB 662 CTCCTTTTGACTTATTTGATGATGAAGATGCTGATATATATCTGATGCTATGCTTTATCG 721  
QY 664 CTTCTTTTATGGGTGGGGGAAAGAGATATACATCCCTTCCTTATGGGATCCAGCT 723  
DB 722 TGCCTCTTCAACGGGATATGGGATTCGGCTTATCTGACACCGTTTACAGATTTCTTCG 781  
QY 724 TACCTGTGCAACCCAGCTCTATGAGCCATCCAGAGATCTTGCTCATATGGGTATGTA 783  
DB 782 CACCTCCACAGCGCATGATCAATGTTCTTCAGAAAGATATTCGGGAATATGTTATGGA 841  
QY 784 GTGTGCTTGGAGAGACGATGAGAAATTTGAGACAGACAGACAGAGAGATGACCA 843  
DB 842 GTGTGCTTGGAGAGACGATGAGAAATTTGAGAGAGACAGACAGAGAGAGATGACCA 901  
QY 844 CGGGGAAATGATGTGTGTGTGATGATGAT-----GACG 876  
DB 902 TCATTTAAGCATGAAGAGAAACAATGTGAGGTTCCATGATGACGACGAACTAGATG 961  
QY 877 ATGCTGATCTACCACTAAATGATGAAGCAAGCAACAACTGTCAGAAATTTCCACTTC 936  
DB 962 ATCTGACATGCTATGATGATGATGAAGAGCAACCTCTCAAGAAAGCTACCTATTTC 1021  
QY 937 CATCAAGCCAGTTAATCCATATAGATGATTAATTAATTCGCTGTGTGTGTGGGT 996  
DB 1022 GTTCAAGCGAATTAATCTTACAGAGATGTTAATCTGTGTGTGCTCGCATTTCTTGTC 1081  
QY 997 TCTTTTCCACACCGAGTGAAGTACCGGTGAAGATGAGATCTTGTGTGTGTGCTCATAT 1056  
DB 1082 TTTCTTTTATTAAGAAATCTCCATCCAGTCAAGTCAATGATGATGATGATGATGATGAT 1141  
QY 1057 CTGTTATCTGTAATCTGTTGTCATGCTTGTGATCTTGTGATCAATTTCCAAAGTGT 1116  
DB 1142 CAGTTATATGCAAGATATGTTGTGAGTGTCTGTGATTTCTGATCAATTTCCCAATGCT 1201  
QY 1117 TCCCTATTGAGAGAGACTTAAGTACCGGCTGTCACTGAGTGTGACAGAGAGGCC 1176  
DB 1202 ATCTCATAGAAAGTAAACATCTCGATGAGCTCTCTCAGTACAGAAAGAGAGAA 1261  
QY 1177 AGCANTCTCACTTCTCCAAATGATTTCTTGTGATGATGATGATGATGATGATGATGAT 1236  
DB 1262 AACCGTCAGATTTACCTCTGTTGATGTTTGTGATGATGATGATGATGATGATGATGAT 1321  
QY 1237 CTCCTTGTGTCACAAATACTGTTCTATCTATCTTGTGATGATGATGATGATGATGAT 1296  
DB 1322 CACCTTGAATTAACAGAAACACAGTTCTTCCATTTAGAGATGATGATGATGATGATGAT 1381  
QY 1297 AGGTTTCTTGTATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
DB 1382 AGGTTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441  
QY 1357 AAACATCGAATTTCCAAAGAAATGAGTCTTCTTGTGCAAGAGTGAATATTTGAACCTTC 1416  
DB 1442 ATACAGCTGAGTTTCTGAAATATGAGTCTTCTTGTGAAAGATTTAATATGAGAGCAC 1501  
QY 1417 GCGCTTCAGAGTGTACTTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1476  
DB 1502 GAGCTCTGAGTGTATTTTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1561  
QY 1477 ACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGATGATGATGATGATGATGATGATGAT 1536  
DB 1562 CTTTGTGTCAGAGAAAGTGTGCTATGAGAGAGATTTAAGAGAGTTTAAAGAGATTA 1621  
QY 1537 ATGCTTGTGTTCCAAAGCCGAGAAAGTCTTCCGAGAGAGAGAGAGATGATGATGATGAT 1596  
DB 1622 ATGCACTGTTCTCTGCAAGAAAGTCTGAGAGAGAGATGATGATGATGATGATGATGAT 1681  
QY 1597 CCGCTGCGCTGAGAAACATGTTCTGATGATCTGATGATGATGATGATGATGATGATGATGAT 1656  
DB 1682 CTCCTTGGCTGAGAAACAGTCCGTGACCAATCTGGAATGATGATGATGATGATGATGATGAT 1741



ORIGIN

Query Match 36.4%; Score 1431.8; DB 6; Length 3255;  
Best Local Similarity 66.0%; Pred. No. 1,6e-266;  
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

QY 144 ATGAGAGCGAGCGCGGGCTGTGGCGGCTCCACAAACCGGAGCGAGCTCGCGTCATC 203  
DB 1 ATGAAATCTGGTGTGGCTCATTTGCTGCTCTACAAACGAAACAAATTCGTTCTCAT 60  
QY 204 CGCCGCGAGCGCGATCCCGGCGGAAACCGCGCGGAGAGAAACCGGAGGTGCCAG 263  
DB 61 AACCGCGATGAGAGTCCGAAATACGATCACTACAAAGAACTGAGTGGCGAAATCTCA 120  
QY 264 ATTTCGCGAGCGAGCTGTGGCTTGCCTCCGCGGAGACCTTTCGTGGCGTCAAGAG 323  
DB 121 ATCTGTGAGATGAATCGAATTCGATTCAGGTTAGCAGTAGAGCTTTTGTGCTTCAACGA 180  
QY 324 TGGGCTTCCCGCTCTGCGGAGCTGTACGAATACGAGCGCGGAGCGGACCGCAAG 383  
DB 181 TGGGCAATTCGGGTTGTAGACCATCTATGAGATGAAAGTAGAAGAAATCAAGCT 240  
QY 384 TGGCCCGAGTGAAGCTGTATCAAGCGCTCAAGGGCTGCCAAGTGTGACGGTAC 443  
DB 241 TGTCTCACTGCAAACTCGATTCAAAGGATTTAAAGTGTCCACGGGTTGATGAGAT 300  
QY 444 GAGAGAGAGAGCGGCTGTATGATCTGCAACGAGTTCACCTGGAGACGGCATGAC 503  
DB 301 GATGAAGAGAGAGATGATGATTCATGATGAGATTTGATCATGGGATGAGCCCT 360  
QY 504 CAGTCTGTGGCGAGTCTCATGCTTACGCGCAATGAGTACGGCGGTGAGAGTCCCT 563  
DB 361 GAACATGCGCGCTAAGACCGGACCTCTTCAAGCTTAAACCGGTGTGTGATTTGAT 420  
QY 564 AATGGCGCGCAAGAGCTTTCAGCTCAACCCCAATGTTCCACTCCCTCAACGGGCAA 623  
DB 421 TCAGCTCACTG-----GCTTCAGATTCCTCTTTGACTTATTTGAT 465  
QY 624 ATGGTGAATGATCCACCGGAGAGCAACGCGTGGTGCCTTCTTTCATGGGTGG 683  
DB 466 GAAGATGCTGATGATGATTCGATTCGATGCTCTTATGCTGCTTCAACGGGATAT 525  
QY 684 GGAAGAGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCGAGTCT 743  
DB 526 GGGAAATCGCTCATCTGCAACGGTTTACAGATTCCTGCACTCCACAGGGAGATCA 585  
QY 744 ATGAGACCATCCAGAGCTTGTGCTCATAGGGAATGATGTTGCTTGGAGAAACG 803  
DB 586 ATGGTCTCTGAAAGATATTTGCGAATATGTTATGAGAGTGTGCTTGGAGAACCT 645  
QY 804 ATGAGAAATTTGAAGAGAGACAAGAGAGATGCAACGAGGGAATGATGGTGGT 863  
DB 646 ATGGAATTTGGAAGAGACGAAAGCGAAGAGCTTCAAGTCAATTAAGTAAAGAGGA 705  
QY 864 GATGATGT-----GACGATCTGATCTACCATATG 896  
DB 706 AACAAATGTCGAGTTCATATGATGACGAACTAGATGATCTGCAATGCTATGATG 765  
QY 897 GAGGAAGCAAGACAACCTGTCCAGAAATTCCTACTTCATCAAGCCAGATTAATCCA 956  
DB 766 GATGAAGGAAGAACCTCTCTCAAGAAAGCTACTATGTTTCAAGAGAAATTAATCT 825  
QY 957 TATAGATGATTAATCATTTATGCGCTGTGTGGGTTTCTTCTTCCACTACCGAGT 1016  
DB 826 TACAGATGTTAATTCGTGTCCCTCGCATTTCTTTCTTTCTTCAATTAAGATTT 885  
QY 1017 ATGATTCGGTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGTAATCTGG 1076  
DB 886 CTCATCTCAGTCAATGATGATGATTTGTTTAACTGATTAATCTCGAGATATGG 945  
QY 1077 TTTGGCATGCTTGGATTTCTGATCAATTCGCAAGGTTCCCTATTTGAGAGAGACT 1136  
DB 946 TTTCAGTGTCTTGATTTGATTCATTCCTCCAAATGTTATCTTAAGAACCTGAAACA 1005

QY 1137 TACCTAGACCGGCTGTCACTAGGTTCCAAAGAAAGCCACGCTCACTGCTCTCA 1196  
DB 1006 TACCTCGATAGACTCTCTCTCAGGTACGAAAGGAAAGAAACCTTCAGGATTAACACT 1065  
QY 1197 ATTTGATTTCTTGTGATGACGTTGATCCCTTAAAGAAACCTCTTGGTGCACAAAT 1256  
DB 1066 GTTGAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1125  
QY 1257 ACTGTTCTATCTATCTTCCGTGATTAATCTGTTGATTAAGTTCTTCTGATGTTCT 1316  
DB 1126 ACAGTTCTTCTCATCTGACAGTTGATTAATCTGTTGATTAAGTTCTTCTGATGTT 1185  
QY 1317 GATGATGTTGCTCAATGCTTAACGTTTGAAGATTAATCTGAAATCTGAAATTTGCA 1376  
DB 1186 GACGATGTTGCTCAATGCTTAACGTTTGAAGATTAATCTGAAATCTGAAATTTGCA 1245  
QY 1377 AATGGGTTCTTCTGCAAGAGTAAATTAATGAACTCGGCTCCAGAGTGTACTTC 1436  
DB 1246 AATGGGTTCTTCTTGTAAAGATTAATTAATGAGCCACGAGCTTCGAGTGTATTTT 1305  
QY 1437 CAACAGAGATGACTACTTGAAGACAAAGTGGCAGCAAACTTGTGTAGGAGAGAGA 1496  
DB 1306 TCTCAGAGATGATTAATCTGAAAGAAACAAAGTTCATCTGCTTTTGTGAGGAGAGT 1365  
QY 1497 GCAATGAGAGAGATGAGGAATTCAGGTGAGATCAATGCTTAAGTTGCCAAAGCC 1556  
DB 1366 CCAATGAGAGAGATTAATGAGAGATTAATGAGATTAATGAGATTAATGAGATTA 1425  
QY 1557 CAGAAATGCTTGAAGAGAGATGAGCAATGCAAGTGAACCCCTGCTGAGAAACAT 1616  
DB 1426 CAGAAAGTCTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1485  
QY 1617 GTTGTGATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1676  
DB 1486 GTCCGTGACCATCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 1545  
QY 1677 GAGGAAATGAACTGCCAGATGTTGTTATGTTTGAAGAGAAACGACAGGCTAATAAC 1736  
DB 1546 GATGTAATGAGATTAACAGCTGATGATGATGATGATGATGATGATGATGATGAT 1605  
QY 1737 CATCATAGAAAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 1796  
DB 1606 CACCAAGAAAGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1665  
QY 1797 GCTCCATTTTGTAACTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1856  
DB 1666 GCTCTTACCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1725  
QY 1857 GAAGCAATGTTTATGATGAGACCTTTTATGAGAAAGAGATTTGCTATGATGATG 1916  
DB 1726 GATCTATGTTGTTTATGATGAGACCGCAATCGGGAAGAAAGTTTGTATGTCAGTT 1785  
QY 1917 CCTCAAGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1976  
DB 1786 CCGCAGAGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1845  
QY 1977 TTTGATATCAATGAGAAAGTTGATGATGATGATGATGATGATGATGATGATGAT 2036  
DB 1846 TTTGATATTAACATGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1905  
QY 2037 TGTGTAATTTGAAGCGAGCATTAATGATGATGATGATGATGATGATGATGATGAT 2096  
DB 1906 TGTGTTTGAAGAGACAGCTCTTATGATGATGATGATGATGATGATGATGATGATG 1965  
QY 2097 TCAAGACTTGAACATCTGCGCCCAAGTGTCTTGTGTTGCTGCTTGTGCAATAG 2156  
DB 1966 GGCAGAAACCTGTAAGTGTGCTTAAATGATGATGATGATGATGATGATGATGAT 2025  
QY 2157 AAGCAAAAGAGACTACCAAAACCAAAACAGAGAGAAAGTATTTATTTTCAAGAA 2216  
DB 2026 AGTAAAG 2073

QY 2217 GAAGAAACCAATCCCTGCATATGCTTTGGTAAATTGACGAAGCTGCTCAGAGCT 2276  
 DB 2074 -----AACCAATTCATGCGCTAGAGATGTCGAGGAAGTGTATGCTCCACAGTGA 2127  
 QY 2277 GAGAAAGAAAGCCGGATTTGTAATCAACAAAATTAGAAAAGAAATTTGGCCATCT 2336  
 DB 2128 AATGTTGAGAAAGATCTGACGACACAAATTGAATTGAGAAAGAAAGTTTGGACATCT 2187  
 QY 2337 TCTGTTTTTGTACATCAACACTTCTGAGAAATGTTGAACCTTGAAAGTGCAGTCT 2396  
 DB 2188 CCGGTTTTGTCCTGCTGCTGTTACAGAAACGGGAGGTTCCCGGTAAACGCAAGCCCC 2247  
 QY 2397 GCTTCTCTTGAAGAAAGCTATACATGTCATTAGTTGGTTAGTAAGAAAGACAGAC 2456  
 DB 2248 GATGTTTGTAAAGAAAGCCATCAAGTTATTAGCTGCGGGATCGAAGATTAACCGAA 2307  
 QY 2457 TGGGAAAAGATGGCTGATCTATGATCGATACAGAAAGATTTCTAACTGGTTTC 2516  
 DB 2308 TGGGAAAAGATCGGGTGAATTTATGATCGGTGACAGAAATATCTCGACGGTTTC 2367  
 QY 2517 AAGATGATTTGTCATGTTGGCGGTCAATTTACTGATACCTTAACGGGTTGCATTCAA 2576  
 DB 2368 AAGATGATTTGTCATGATGAGATCTGATGATGCTTAAGCGGTGACGCTTTTAA 2427  
 QY 2577 GGTTCGACCTCTGAATCTTTCAAGATCGTCTTCAACAGGTGCTTCGGGGCTCTGGG 2636  
 DB 2428 GATCTGCTCTTAACTTGTGATGATGCTTCAAGTTCTTAAGTTGGGCTCTGGC 2487  
 QY 2637 TCTAATGATGATCTTCTTCAAGATCAATGCCCCCTTTGGTATGGATGGGCGTCTG 2696  
 DB 2488 TCTGTAAGATTTTCTTGAAGACATTTGCGATATGATGATGATGATGATGATGATGAT 2547  
 QY 2697 AAATTTTGGAAAGATTTTCTTCAATCAATCCATGATGATGATGATGATGATGATGAT 2756  
 DB 2548 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2607  
 QY 2757 CTTTGGCTTATCTGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2816  
 DB 2608 TTGATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2667  
 QY 2817 GAGCTGAATATGTTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2876  
 DB 2668 GAGATGAGACATGACGAGATATCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2727  
 QY 2877 AGCATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2936  
 DB 2728 GGAATCTGAAATGCAATGCGGAGGTGTCGAAATGATGATGATGATGATGATGATGATGAT 2787  
 QY 2937 TTTCTGGTATGAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2996  
 DB 2788 TTTTGGGTAATCGAGGGGCTCTCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2847  
 QY 2997 GTCATGAGTGTGATGATCAAGCTCACCGTACATCAAGGGGAGAGTGAAGAG 3056  
 DB 2848 GTTCTAGCGGAGTTAAACGAATTTCAAGTCACTTCAAAAGACAGACATGAGCT 2907  
 QY 3057 TTCTGAGCTATATACATTTCAATGAGTCACTTATGATGATGATGATGATGATGATGATGAT 3116  
 DB 2908 TTCTGAGCTTAACTTTCAAGTGAACACTTGTGATGATGATGATGATGATGATGATGATGAT 2967  
 QY 3117 CTAATGAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3176  
 DB 2968 ATCATTAATCATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3027  
 QY 3177 TCAATGAGGAGGCTCTTGGGAGCTATCTTGGATTTGGTGTGATGATGATGATGATGATGAT 3236  
 DB 3028 TCAATGAGGAGGCTCTTGGGAGCTATCTTGGATTTGGTGTGATGATGATGATGATGATGAT 3087  
 QY 3237 CCTTTCTGAAGGTTGTTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3296  
 DB 3088 CCAATCTCAAGGAGATGCTTGGAG 3147  
 QY 3297 TCCATTTCTGCTGCTCAATCTTCTGCTCTTGGGTTGAGATGATGATGATGATGATGATGAT 3356

DB 3148 TCTATTTCTTACGCTTTCATGATCTTGTGGGTCAAGATTAACCGTTGTGGCT 3207  
 QY 3357 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3397  
 DB 3208 A---AAGGGGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3245  
 RESULT 9  
 AR267558  
 LOCUS AR267558 3828 bp DNA linear PAT 10-APR-2003  
 DEFINITION Sequence 7 from patent US 6495740.  
 ACCESSION AR267558  
 VERSION AR267558.1 GI:29697634  
 KEYWORDS  
 SOURCE  
 ORGANISM Unknown.  
 REFERENCE  
 1 (bases 1 to 3828)  
 AUTHORS Arioli, A., Williamson, R.E., Betzner, A.S. and Peng, L.  
 TITLE Manipulation of cellulose and/or .beta.-1,4-Glucan  
 JOURNAL Patent: US 6495740-A 7 17-DEC-2002;  
 FEATURES  
 source  
 1..3828  
 /organism="unknown"  
 BASE COUNT 1056 a 721 c 912 g 1139 t  
 ORIGIN  
 Query Match 36.3%; Score 1429.4; DB 6; Length 3828;  
 Best Local Similarity 65.8%; Pred. No. 4,8e-266;  
 Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;  
 QY 124 GAATGAGAGGAGAGAGAGATGAGAGGCGGAGCTGTGGCCGGCTCCCAAC 183  
 DB 219 GATTCGGTGTAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278  
 QY 184 GCAAGACCTGCTGCTGATCCCGCGAGCGAGATCCCGGCGGAGCCGCGGAGC 243  
 DB 279 GAAAGAAATTCGTTCTCATTAACGCGATGAGAGGCGAATACGATCAATGATCAAGAAC 338  
 QY 244 AGAAGGAGAGGATGTCAGATTTGCGGAGAGAGCTGCGCTTCCCGCGGAGAGC 303  
 DB 339 TGAAGGCGCAACATGTCATATCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 398  
 QY 304 CTTTGTGGCGTCAAGAGATGCGCTTCCCGTGTGCGGAGCTGCTAAGATACGAGC 363  
 DB 399 TCTTTGTGCTTGAACAGAAATGCGATTCGCGTGTGTAAGACATGATGATGATGATGATGAT 458  
 QY 364 GCGGAGAGGAGCGAGAACTGCCCACTGCAATCTGATACAAAGCGCTCAAGGCT 423  
 DB 459 GTAGAGAAAGAAATCAACTGCTGCTCACTGATGCAAACTGATACAAAGGATTAAGATA 518  
 QY 424 GCCAAGTGTGACCGGTGACGAG 483  
 DB 519 GTTCAAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578  
 QY 484 ACTGGAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543  
 DB 579 TTGATCATGAGATGAGACCTGTAACATGCGCTGAAGCGCATCTTCAAGCCTTAACA 638  
 QY 544 ACGGCGGTGAGAGTACCTTAATGCGCGGCAAGCTTCCAGTCAACCCCAATGTTT 603  
 DB 639 CCGGTCGTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683  
 QY 604 CACTCTCTCAACAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
 DB 684 CTTTCTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 743  
 QY 664 CTTCTTCAATGAGTGTGGGAGAGAGATACATCCCTTCTTATGAGATCCAGCT 723  
 DB 744 TCCCTCTTCAACGGGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 803  
 QY 724 TACCTGTGACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783



Db 804 CACCTCCAGCGGAGATCATAGTCTCTCAGAAAAGATATGGGAATATGTTATGAA 863  
Qy 784 GTGTTCCTTGAAGAAAGAGATGAGAAATTTGAAGCAGAGCAAGAGAGATCAGCAGA 843  
Db 864 GTGTTCCTTGAAGAAAGAGATGAGAAATTTGAAGCAGAGCAAGAGAGATCAGCAGA 923  
Qy 844 CCGGGAATGATGCTGCTGATGATGCT-----GACG 876  
Db 924 TCATTAAGCATGGAAGGAGAAACATGCTGAGTTCCATGATGACGACGAATGATG 983  
Qy 877 ATGCTGATCTACCACTAATGATGAGCAAGACAAACATGCTCAGGAAATTTCCACTTC 936  
Db 984 ATCTGACATGCTTATGATGATGAGAAAGAAACATCTCTCAGAAAGACTATATTC 1043  
Qy 937 CATCAAGCCAGATTAATCCATATAGATGATATATCATATATTCGCTTTGTGTTGAGG 996  
Db 1044 GTTCAAGCAAGAAATTAATCTTACAGATGTTAATTCGTGCTGCTGCGATTCCTGCTG 1103  
Qy 997 TCTTCTTCCATACCGAGATGATCGGATGATGATGATGATGATGATGATGATGATGAT 1056  
Db 1104 TTTCTTTTCATTAAGATATCTCAGTCAGTCATGATGATGATGATGATGATGATGAT 1163  
Qy 1057 CTGTTATCTGTGAATCTGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116  
Db 1164 CAGTTATATGCGAAATATGCTTTCAGATGCTTTCAGATCTTTCAGATCTTTCAGATCT 1223  
Qy 1117 TCCCTATTGAGAGAGACTTACCTAGACCGGCTGCTCAGAGTTGCAAGAAAGGCC 1176  
Db 1224 ATCTTATAGAGAGAGAAACATACCTGATGATGATGATGATGATGATGATGATGATGAT 1283  
Qy 1177 AGCAGCTCACTTCTCTCAATGATTTCTTTGTCAATGATGATGATGATGATGATGATGAT 1236  
Db 1284 AACCGTCAGAGATGAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1343  
Qy 1237 CTGCTTTGTGCAACAAATCTGCTCTATCTATCTCTTGGGATGATGATGATGATGATGAT 1296  
Db 1344 CCCCCTGATTAAGCAACAGATCTTTCATCTCAGAGATGATGATGATGATGATGATGAT 1403  
Qy 1297 AGGTTTCTGCTATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
Db 1404 AGGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
Qy 1357 AAACATCTGATTTGCAAGAAATGCTTCTTTCGCAACCGTACATATTTGAACCTC 1416  
Db 1464 ATACAGTGTATTTGCTACAAATGCTTCTTTCGCAACCGTACATATTTGAACCTC 1523  
Qy 1417 GCGCTCAGAGTGTCTCTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1476  
Db 1524 GAGCTCTGAGTGTATTTTCTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1583  
Qy 1477 ACTTTTGAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1536  
Db 1584 CTTTGTGAGGAGAAAGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1643  
Qy 1537 ATGCTTATGTTCCAAAGCCGAGAAATCTTCTGAAGAGATGATGATGATGATGATGATGAT 1596  
Db 1644 ATCAGCTGTGTCTACGCAAGAAATGCTGAGAACTTGAAGATGATGATGATGATGATGAT 1703  
Qy 1597 CCCCCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656  
Db 1704 CTTCTTGGCTGAGAAACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1763  
Qy 1657 AAAGCGAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1716  
Db 1764 ATGATGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1823  
Qy 1717 AGAAAGCAGAGCTATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1776  
Db 1824 AGAAAGCAGAGCTATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883  
Qy 1777 TCTGCTGATCTATCAAAATGCTCATATTTGTTAACTTGAATGATGATGATGATGATGATGAT 1836

Db 1884 TCTGCTGATCTATCAAAAGCTCTTACTTCTTAATGATGATGATGATGATGATGATGATGAT 1943  
Qy 1837 ACAACGAAAGCTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1896  
Db 1944 ACAACGAAAGCTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003  
Qy 1897 AGGTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1956  
Db 2004 AAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063  
Qy 1957 CTAAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2016  
Db 2064 CAACCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2123  
Qy 2017 CAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2076  
Db 2124 CATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2183  
Qy 2077 CCAAAAGAAAG 2136  
Db 2184 CAAG 2243  
Qy 2137 GTTGTGCTTTGCAATGAG 2196  
Db 2244 GTTGTGCTTTGCAATGAG 2303  
Qy 2197 AGTTATATTTTCAAGAAAG 2256  
Db 2304 AGACTTCA-----AGCAGATTCATGCTGCTAGAGAAATGCTGAGAG 2345  
Qy 2257 AGGAGCTCTCAG 2316  
Db 2346 GTGTTATCTCCAG 2405  
Qy 2317 AAAAGAAATTTGAG 2376  
Db 2406 AG 2465  
Qy 2377 CTTGAG 2436  
Db 2466 TTCCCGGTAAG 2525  
Qy 2437 GTTATGAG 2496  
Db 2526 GGTACCAAGATTAACCGAGATGAG 2585  
Qy 2497 AAGATTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2556  
Db 2586 AAGATTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2645  
Qy 2557 CTAAACGAGTTCATTAAG 2616  
Db 2646 CTAAAGCTCAGCTTTTAAAG 2705  
Qy 2617 TGTTCGAG 2676  
Db 2706 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2765  
Qy 2677 ATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2736  
Db 2766 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2825  
Qy 2737 ATCTTGAATCTATTCCTCTTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2796  
Db 2826 ATCTTGAATCTATTCCTCTTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2855  
Qy 2797 CAGGAAATTTATCACTCCAG 2856  
Db 2886 CAGGAAATTTATCACTCCAG 2945  
Qy 2857 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2916  
Db 2946 TCATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3005

QY 2917 ACTGGTGGAGGATGAGCATTCTGGTCAATTGAGAGTGTGCTCCACACTCTTTGCTG 2976  
 DB 3006 ATGTGTGGAGAAACGAGCAGTTTGGTAAATCGAGAGGCGCTCCGCAATCAATTGCTC 3065  
 QY 2977 TGTTCACGAGGACTTCAAGTCATCATCTGAGTGTGATCAAGCTTCAACCGTGAACATCA 3036  
 DB 3066 TGTTCACGAGGTTTCTCAAGTCTTCTGAGCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 3125  
 QY 3037 AGGTGGAGATGATGAGAGTCTCAAGCTTATATCAATTCATCAATGAGTCTTCAATGTA 3096  
 DB 3126 AGCAGCAGACATGAGAGCTTCTGAGCTTATCAATTCATCAAGTCTTCAAGTCTTCA 3185  
 QY 3097 TACCTCTCAACACTTCTTCTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 3156  
 DB 3186 TTTCTCCGCAACACTCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3245  
 QY 3157 GCATCAATAGAGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3216  
 DB 3246 CCAATTCAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3305  
 QY 3217 GGTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3276  
 DB 3306 GGTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3365  
 QY 3277 CAACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3336  
 DB 3366 CTACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3425  
 QY 3337 GGTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3396  
 DB 3426 GATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3482  
 QY 3397 G 3397  
 DB 3483 G 3483  
 RESULT 10  
 AX030944 3828 bp DNA linear PAT 24-NOV-2000  
 LOCUS Sequence 7 from Patent WO9800549.  
 DEFINITION AX030944 AX030958  
 ACCESSION AX030944.1 GI:10278347  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1  
 Williamson, R.E., Peng, L., Arioli, A. and Betzner, A.S.  
 Manipulation of cellulose and/or beta -1,4-glucan  
 Patent: WO 9800549-A 7 08-JAN-1998;  
 WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI  
 ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;  
 COMMUN SCIENT IND RES ORG (AU)  
 On Oct 15, 2002 this sequence version replaced gi:10278361.  
 COMMENT  
 FEATURES  
 source  
 1..3828  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="COLUMBIA"  
 /db\_xref="taxon:3702"  
 /clone="ATH-A"  
 239..3493  
 /note="unlabeled protein product"  
 /codon\_start=1  
 /protein\_id="CAC09626.1"  
 /db\_xref="GI:10278348"  
 /db\_xref="REMBL:CAC09626"  
 /translation="MNTGRLIAGSHNRPFVLIINAEASRIISVQELSGCTCOTCGD  
 ELETVSSELPVACNECAPVPCPTCYERREBONACPOCKTRIKRIKSPVDDGDDE

EEDIDLEVEFDHGMDEPHEAAEALSSRLNTRGRGLDAPPGSOIPLITTCDEBADM  
 YSDRALIVPSTGCMNVVAPPTDSSAPPOASMPVOKOIAETGVSVAMKRMV  
 MKRQGEKLVQVIRKRGNGSGDDDELDPDPMDEBQPLSRKPISSSRINPY  
 RMLILCRALIGLEFPHYRIHPVNDAYGLMTSVICETIWFVSLIDFPKMYTIRE  
 IYLRPLSRVREKEKPSGLAPVDVDFVSLDLKBPILITANTVLSILADVDPVAVAC  
 VYSNNGAMLTPEALSDPADATKVPCEKFNEMPRAPYVSGSKMVLNKKVPAF  
 VREBRAMKRDVEEPFKVKNALVATPAOKREBRMTMODCTPMPGNVVRHPKQIVFLG  
 HSGARDNDGNLPLVYVSRERKRGFDHKKKAGAMNLSIRISATVLSNAPYLLNDCD  
 YINNSKATRESKCFMMDPQSGKVCYVQFPQFDGIDHNDYISNRYVFDINMGD  
 GIQEPYVGTCCVPERKQALYGFDAPEKPKPKPKCCKPCKCCLCCGLRKKSKTATD  
 KKTINPESKOIHALENDGIVPVSVNREKSEATQLEKPKFQSGSVYVAVYALN  
 GGVPRNAPACILREAIQVISCQVNTKEWGEIEMIGVSTVEDLITFPMKCHMRS  
 VYCMKRAFPKGSAPINLSRLHOVLRAKLSVEIFLSRHQPIWYGVGGIKWIFRS  
 YINSVYVWTSPLPIVYCSLPAVCLIRKPYEPESNAYGLIFMLMFSIAVTGLEM  
 OMGVGIDMWRNEQVWIGASLSHLPFLQGLKVLGVNTFTVTSKADDAFSE  
 LYIFKMTLLIPPTLLIINITIGVIVGSDAISNGYDMPGLFGLFALWVYVLYP  
 FLKMLGKQDMPTLIIVWSILSLILTLWVRINPFAKXGPLYEIGLNCGN"  
 BASE COUNT 1056 a 721 c 912 g 1139 t  
 ORIGIN  
 Query Match 36.3%; Score 1429.4; DB 6; Length 3828;  
 Best Local Similarity 65.8%; Pred. No. 4.8e-266;  
 Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;  
 QY 124 GAAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 183  
 DB 219 GATTCGGTGTGAAACATCATGAACTGATGATGATGATGATGATGATGATGATGATGAT 278  
 QY 184 GCACGAGCTGCTGCTCATCCGCGCGAGGAGATCCGCGCGAGGAGGAGGAGGAGGAGGAGG 243  
 DB 279 GAAACGAAATTCGTTCTTCAATTAAGCGGATGAGAGAGTCCGAAATAGATCATGACAAAG 338  
 QY 244 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303  
 DB 339 TGAGTGGGCAAAATGATCAATCTGTGAGATGAAATGAAATGAAATGAAATGAAATGAA 398  
 QY 304 CTTGCTGGCGTGCACAGAGTGGCGGCTTCCCGCTGCGCGGAGCTGCTCAAGATACAGAG 363  
 DB 399 TCTTGTGCTTGCACAGAGATGCGATTCGCGGTTGAGACCAATGCTATGATGATGATGATG 458  
 QY 364 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 423  
 DB 459 GTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 518  
 QY 424 GCCAAGTGTGACCGGTGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483  
 DB 519 GTCCACGGGTGATGAGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 578  
 QY 484 ACTGGAGGCGCATGATGAGCTGTGTGCGCGGAGTCCATGCTTCAAGCCCAATGAGCT 543  
 DB 579 TTGATCATGAGATGAGACCTGTAACATGCGCGTGAAGCCGCACTCTTCAACGCTTAA 638  
 QY 544 ACGGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 603  
 DB 639 CCGGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 683  
 QY 604 CACTCTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 663  
 DB 684 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743  
 QY 664 CTTCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723  
 DB 744 TGCCTCTTCAACGAGGATATGAGGATATGAGGATATGAGGATATGAGGATATGAGGAT 803  
 QY 724 TACCTGTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783  
 DB 804 CACTTCCACAGGCGAGATCATGATTCCTCAAGAGATTTCCGAAATATGATGATGATGAA 863  
 QY 784 GTTGTCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 843  
 DB 864 GTTGTCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 923  
 QY 844 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 876

Db 924 TCATTAGCATGAGGAGAAACAATGCTGAGGTTCCAAATGATGACGACGAACCTAGATG 933  
 Qy 877 ATGCTGATACCACTAATGATGATGAAGCAAGACAAACAATGTCAGAGAAAATTCACCTG 936  
 Db 984 ATCTGACATGCTTAATGATGATGAGAGAGAAACAACCTCTCTAAGAAAGCTACATATTC 1043  
 Qy 937 CATCAAGCCAGATTAATCCATATAGATGATATCATTAATTCGGCTTGTTGGGGT 936  
 Db 1044 GTTCAAGCAAGATAAATCCTTACGAGATGTTAATTCGTGTCGCTCGCATTCCTGGTC 1103  
 Qy 997 TCTCTTCCACTACCGATGATGATCCGGTGAATGATGATTTGCTTTGGCTCATAT 1056  
 Db 1104 TTTTCTTCAATTAAGAAATCTCANTCAGTCAATGATGATGATTAATGATTAACGT 1163  
 Qy 1057 CTGTTATCTGAAATCTGTTGGCCATGCTTGATTTCTTGATTCATATCCCAAGAGGT 1116  
 Db 1164 CAGTTATATCGAAATATGTTGTTGCAAGTCTTGATTTCTTGATTCATATCCCAATGCT 1223  
 Qy 1117 TCCCTATAGAGAGAGACTTAACCTAGACCCGCTGTCTCAGAGGTTGCAAGAAAGCC 1176  
 Db 1224 ATCTATAGAAACGTGAACAATACCTGATAGACTCTCTCAGGTACGAAGAAAGAA 1283  
 Qy 1177 AGCCATCTCACTTGTCTCAATGATTTCTTGTCACTACGTTGATCCCTTAAAGAAC 1236  
 Db 1284 AACCGTCAGGATTAGACCTGTGATGTTTGTATGATACAGTGAATCCGTTGAAGAGC 1343  
 Qy 1237 CTCCTTGGTACAAACAATACCTGTTATCTTATCTTGGTGATATCTGTTGATA 1296  
 Db 1344 CCCCCTGATTTACAGAAACAGATCTTTCATCTTACAGATTTGATTAATCTGTTGATA 1403  
 Qy 1297 AGGTTCTGCTATGTTTCTGATGATGCTGCAATGCTAAGCTTGAAGCATTAATCTG 1356  
 Db 1404 AGGTGCTGTTATGATTAACAACAATGCTGAGCTATGCTTACATTTGAAGCTCTCTG 1463  
 Qy 1357 AAACATGTGAATTTGCAAGAAATGGTCTTCTTCTGCAACGTTACATATTTGAACCTC 1416  
 Db 1464 ATACAGCTGATTTTGTACAAAAATGGGTTCTTTTGTAAAGATTAAATGCAAGCAC 1523  
 Qy 1417 GCGCTCAGAGTGTACTTCCAAACAGAAATAGACTCTTGAAGAAACAAGTGGACAA 1476  
 Db 1524 GAGCTCTGAGTGTATTTTCTCAGAAATGATTTACCTGAGAAACAAGTTCACTCTG 1583  
 Qy 1477 ACTTGTAGAGAGAGAGACATAGAAAGAGATAGGAATTCAGAGTGAAGATCA 1536  
 Db 1584 CTTTGTGACGGAACGTGCTATGAAAGAGATTAATGAAGATTAAAGTGAAGTAA 1643  
 Qy 1537 ATGCTTAGTTGCCAAAGCCAGAAAGTTCTGAAAGAGTGAACAATGCAAGATGAA 1596  
 Db 1644 ATGCACTGTTGCTACTGCAAGAAAGTCTGAGAAACGTTGACTATGCAAGATGAA 1703  
 Qy 1597 CCCCCTGAGGAAACAATGTTGTATCATCTGGAATGATTCAGGTCTTCTGGCC 1656  
 Db 1704 CTCTTGTGCTGAAACAACGTCCGTGACATCTGGAATGATTCAGGTCTTCTGGCTC 1763  
 Qy 1657 AAAGCGAGAGCTTGACTGTGAGAGAAATGAACCTGCCAGATTTGTTATGTTCTAG 1716  
 Db 1764 ATAGTGAAGTTGCGATAGCGATGTATGATGATACACAGCTATGATGATTTCTGCTG 1823  
 Qy 1717 AGAAAGACAGGCTATTAACCATATTAAGAAAGCTGCTGATGATGATGCTGCTGAG 1776  
 Db 1824 AGAAGCGGCTGGAATTTGATACCAAGAAAGCTGAGGCTATTAATCTTGTGATCCGAG 1883  
 Qy 1777 TCTGTGCTACTAACAATGCTCATATTTGTTAACTTGGAATGTTGATACATCAATCA 1836  
 Db 1884 TCTGTGCTTCTATCAAAAGCTCTTACCTTTAATGTCGATGATGATCCTACTACATCA 1943  
 Qy 1837 ACAACAGCAAGGCTATAAGAGAACATGTGTTTATGATGACCCCTTACTAGAAAGA 1896  
 Db 1944 ACAACAGCAAGGCAATTAAGAAATCTATGTTTCAATGATGACCCGGAATCCGGAAGA 2003  
 Qy 1897 AGGTTTGTATGTAAGTTCCTCAAGATTTGATGGATTTGATCGCATGACCGATATG 1956

Db 2004 AAGTTGTTATGTTCAAGTTCCGCAAGATTTGATGGCATTTATGACATGATTAATCT 2063  
 Qy 1957 CTAAACCGAATGTTGCTTTTGTATATCAATGAAGAGTTGATGTTATTCAGGGTC 2016  
 Db 2064 CAACCGTAAACGTTGTTCTTTGATATTAATCATGAAGAGTCTTGATGGATACAAAGAC 2123  
 Qy 2017 CAATTTATGTTGCTACGATGTTATTTAAGAGGAGCATTAATATGTTATGATGCC 2076  
 Db 2124 CGATATATGCTGGAGACAGGTTGTGTTAGAAACAGGCTTTATGTTGTTGATGAC 2183  
 Qy 2077 CCAAAACAAGAGCCACCATCAAGACTTGCACTGCTGGCCCAAGGTGCTTTGCT 2136  
 Db 2184 CAAGAGAGAGAAACCAACAGGCAAACTGTATCTGTTGCTTAATATGTTGTTGT 2243  
 Qy 2137 GTTGCTCTTTGGCAATAGAAAGCAAAAGAACATCAACAAACCAAGAGAGAA 2196  
 Db 2244 GTTGCTGTTGAAGAAAGAGATTAACCAAGAACCAATTAAGAAACTTAACCTAAAG 2303  
 Qy 2197 AGTTATATTTTCAAGAAAGAGAACCAATCCCTGCAATGCTCTTGATGAATG 2256  
 Db 2304 AGACTTCA-----AAGCAGATTCATGCTGAGAAATGTCAGAG 2345  
 Qy 2257 AGAAGCTGCTCAGAGGCTGAGAAATGAAGCCGCTTTGTAATCAACAAAAATG 2316  
 Db 2346 GTTATATGTTCCAGGTCAATGTTGAGAGAGATCTAAGCAACAAATGAATGG 2405  
 Qy 2317 AAAAGAAATTTGGCAATCTCTGTTTGTATCACTCACTCTCGAAATGATGAA 2376  
 Db 2406 AGAAGAGTTTGGACAAATCTCGGTTTGTGCTCTGCTCTGTTCTACAGAACGTTGAG 2465  
 Qy 2377 CCTGAAGAGTCAAGTCTGCTCTCTTTGAAGAGCTATACATGCTATTAAGTTG 2436  
 Db 2466 TTCCCGTAAACGACACCCCGCATGTTGTTAAGAGAACCATTCAGTTAATGAGCTGG 2525  
 Qy 2437 GTTATGAACAAAGACAGCTGGGAAAGAGATTGGCTGATCTATGATGATGATGAC 2496  
 Db 2526 GGTACCAAGATTAACCGAATGGAGAAAGATCGGGATTAATGATGATGATGATGAC 2585  
 Qy 2497 AAGATATTTAACTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATG 2556  
 Db 2586 AAGATATCTGACGAGGTTTCAAGATGATGATGATGATGATGATGATGATGATG 2645  
 Qy 2557 CTAAACGGTGTGATTAAGATTTCTGCACTCTGCAATCTTTCAATGCTCTTACCAAG 2616  
 Db 2646 CTAAAGCTGAGCTTTTAAAGATCTGCTCTATTAATGTCAGATGCTCTTCAACAG 2705  
 Qy 2617 TGTGCTGGGAGCTCTGGGCTATGAGATCTCTTCAAGATCAATTCCTCTTGTG 2676  
 Db 2706 TTTAGCTTGGCTCTTGGCTCTGTAAGATTTTCTGAGACAGCAATTTGCGATATGG 2765  
 Qy 2677 ATGGTATGAGTGGGCTGAAATTTTGAAGATTTTCTTACATCAATCCATGCTGT 2736  
 Db 2766 ATGTTATGATGATGATTAAGATGTTGTAAGAGATTTCTTACATCAATCTGCTGT 2825  
 Qy 2737 ATCTTGAACATTAATCCCTCTTGGCTTACTGATCAATTTGCTGCAATCTGTTATGA 2796  
 Db 2826 ATCTTGAACATTAATCCCTCTTGGCTTACTGATCAATTTGCTGCAATCTGTTATGA 2885  
 Qy 2797 CAGGAAATTTATCACTCAGAGCTGAATTAATTTGCAAGCTGCTGCTGCTGCTGCT 2856  
 Db 2886 CAGGAAATTTATCACTCAGAGCTGAATTAATTTGCAAGCTGCTGCTGCTGCTGCT 2945  
 Qy 2857 TTATGCAATTTTGTACAGAGCTAGAAATGATGATGATGATGATGATGATGATG 2916  
 Db 2946 TCAATATCAATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 3005  
 Qy 2917 ACTGTTGAGAGATGAGCAATTTGATGATGATGATGATGATGATGATGATGATG 2976  
 Db 3006 ATTGTTGAGAGATGAGCAATTTGATGATGATGATGATGATGATGATGATGATG 3065  
 Qy 2977 TGTTCAGGAGCTTCAAGGCTATGCTGATGATGATGATGATGATGATGATGATG 3036  
 Db 3066 TGTTCAGGCTTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 3125

Oy	3037	AGCGTGGAGTATGATAGAGAGTTTCAGAGCTATATACACTTGAATGACTACCTTATGGA	30368
Db	3126	AAGCAGCAGACGATGAGACCTTCTCTGAGCTTTATACATCTTCAAGGACCAACTTGTGGGA	31883
Oy	3097	TACCTCCACACACCTTGCTCTTATGAACTTCATTTGGTGTGCTGCGCTTCAATG	31566
Db	3186	TTCTCTCCGACCAACTTCTGATCATTTAACATCATTTGGAGTTATGTGCGGCTTCTGATG	32454
Oy	3157	CGATCAATPAACGATATGATGATGAGGAGGAGCCCTCTTTGGAGAGCTATTTCTTTGCATTTT	32168
Db	3246	CCATTAGCAAAAGGCTATGACTATGATGAGGAGACCTCTTTGGAGAGACTTTTCTTCGCTCTTT	33059
Oy	3217	GGGTGATTTGTCATCTTATATCCCTTCTTCAAAAGTTTGGTTGGAGAAACAGAGAAC	32767
Oy	3277	CAACGATTTGATCGTCTGTCATCTTCTGCTGAGCTTCAATCTTTCGCTCTTTGGGTTTC	33368
Db	3366	CTACGATTTATGTGTGCTGTGCTATCTTTCTAGCTTCATCTTGACATCTCTTGTGGCTGA	34253
Oy	3337	GGATTGATCTTTCTCTTGCGAAGATGATGTCGCTCTTGAGAGATGTGGTTTGATTT	33968
Db	3426	GAATTAACCCGTTTGTGGCTA--AAGGGGGACGATGTTGGAGATCTGTGCTGTAATTT	34828
Oy	3397	G 3397	
Db	3483	G 3483	
RESULT 11			
BD022677			
LOCUS	BD022677	3828 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Manipulation of cellulose and/or beta-1,4-glucan.		
ACCESSION	BD022677.1	GI:22563900	
VERSION	JP 2001510326-A/5.		
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	Ariori,A., Williamson,R.E., Bettensar,A.S. and Penn,L.		
AUTHORS	Manipulation of cellulose and/or beta-1,4-glucan		
TITLE	Patent: JP 200150326-A 5 31-JUL-2001;		
JOURNAL	THE AUSTRALIAN NATIONAL UNIVERSITY, COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION		
COMMENT	OS Arabidopsis thaliana (thale cress)		
	PN JP 200150326-A/5		
	PD 31-JUL-2001		
	PF 24-JUN-1997 JP 1998503665		
	PR 27-JUN-1996 AU PO 0699		
	PI ANTONIO ARIORI, RICHARD EDWARD WILLIAMSON, ANDREAS STEPHAN PI BETTSUNAR,		
	PI LYONKAI PENN		
	PC C12N15/09,A01H5/00,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/		
	PC 10,C12N5/10,		
	PC C12N9/10,C12N15/00,C12N5/00,C12N5/00		
	CC Strandedness: Single;		
	CC Topology: Linear;		
	CC Key		
	FT Location/Qualifiers		
	FT CDS	239..3490.	
FEATURES			
source	Location/Qualifiers		
	1..3828		
	/organism="Arabidopsis thaliana"		
	/mol_type="Genomic DNA"		
	/db_xref="taxon:3702"		
BASE COUNT	1056 a 721 c 912 g 1139 t		
ORIGIN			
Query Match	36.3%; Score 1429.4; DB 6; Length 3828;		
Best Local Similarity	65.8%; Pred.No. 4.8e-266;		

[illegible]

Qy	1177	AGCCATCTCAACTTGTCTCCAAATTTGATTTCTTTGTGCACTACGGTGTGATCCCTTAAAGAAC	1236
Db	1284	AACCGTCAGGATTAGCAACCTGGTGAATGTTTTTGTATATACAGTGAATCCGTTAAAGAGC	1343
Qy	1237	CTCCCTTGGTGCACAACAAAATACTGTTCTATCTATCCTTTGGTGAATATCTGTGATTA	1296
Db	1344	CCCCCTTGATTACAGCAAAACAGACTCTCTTCATCTTACAGAGTTGATTAATCTGTGGATA	1403
Qy	1297	AGGTTTCTTGCTATGTTTCTGATGATGCTGTGCTCAATGCTAACGTTTGAAGCATATCTG	1356
Db	1404	AGGTTGGGTGTTATGATATCAAAACAAATGTGACAGTAACTTACATTTGAAGCTCTCTG	1463
Qy	1357	AAACATCTGAATTTGGCAAGAAATGGGTCCTTTCTGCAAAACGGTACAAATATGAACTTC	1416
Db	1464	ATACAGCTGAATTTTGTCTACAAATATGGTTCCTTTTGTAAAGAGTTTAAATATGAGCAC	1523
Qy	1417	GCCTCTCAGAGTGTACTTCCMACAGAAATGACCTTGAAGACAGAGTGGCAGCAA	1476
Db	1524	GAGCTCCGATGGTATATTTTCTCAGAAAGATGATTAACCTGAAGAACAAAGTTCAATCTG	1583
Qy	1477	ACTTTGTTAAGGAGAGAGACAAATGAAAGAGATATGAGGAATTCAGTGTAGATCA	1536
Db	1584	CTTTTGTCAAGGAAACGTCTGTCTATGAAAGAGATTAATGAAGATTTTAAAGTAAAGTAA	1643
Qy	1537	ATGCTCTTAGTTCCCAAAAGCCCAAGAAAGTTCCTGTAAGAAAGATGACAACTCAAGATGAA	1596
Db	1644	ATGACACTGTGTTGCTACTACGACACAGAAATGCTCTGAGAAACGTTGGAATCTATCCAAATGAA	1703
Qy	1597	CCCCCTGGCCCTGAAACAAATGTTGCTGATCATCCTGGAATGATTCAGGTCTTCTTGACC	1656
Db	1704	CTCTCTGGCCCTGAAACAAACGTCCTGACCATCTCTGAAATGATTCAGGTGTTCTTGAGTC	1763
Qy	1657	AAACCGGAGGCTTGAACGTGAGAGGAAATGAACTGCCACAACTTGGTTATGTTCTTGAG	1716
Db	1764	ATAGTGGATTCGTGATACGAGATGATTAATGATTAACACGTTATGTTATGTTTCTGTG	1823
Qy	1717	AGAAACGACACAGGCTATTAACCATATAGAAAGCTGTGCTTAATGATTCATTTGCTCCGAG	1776
Db	1824	AGAAAGCGGCCGTGATTTGATCACCACAAAGAAAGCTGAGCTATGAATTCCTGTATCCGAG	1883
Qy	1777	TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAACTTGGATTTGGATCACTACATCA	1836
Db	1884	TCTCTGCTGTCTATCAAAACGCTCCTTAACCTTCTTAATGTGATTTGATCACTACATCA	1943
Qy	1837	ACAACAGCAAGGCTATTAAGAGAACCAATGTGTTTATATGATGAGACCTTTACTAGGAAGA	1896
Db	1944	ACAACAGCAAGCAATTTAGAGAAATCTATGTGTTTCAATGATGAGACCCGCAATCCGGGAAGA	2003
Qy	1897	AGGTTTGCTATGTACAGTTCCTCAAAAGATTTGATGGATGTATGCCATGACCGATATG	1956
Db	2004	AAGTTTGTATGTTCAAGTTTCCGACAGATTTGATGGATGTAATGACATGATGATATCT	2063
Qy	1957	CTAACCCGGAATGTGTCCTTTTATATCAATCAAGTAAAGGTTTGGATGATTAAGGGTC	2016
Db	2064	CAAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAGAGCTTGATGGATACAAAGAC	2123
Qy	2017	CAATTTATGTTGTAAGTATGATGTGATTTTAGAAGCAGCAGCTTATATGTTATGATGCC	2076
Db	2124	CGATATATGTGGGAGCAGGTGTGTGTTTAGAAACAGGCTCTTATATGTTTATGATCAC	2183
Qy	2077	CCAAACCAAGAGACCATATCAGAGACTTGCAACTGCTGGCCCAAGTGGTCTTTTGCT	2136
Db	2184	CAAAACAAAGAAACACACAGGCAAAACCTGTAACTGTGTGACCTTAAATGAGTGTGTTGT	2243
Qy	2137	GTTGCTGCTTTGGCAATAGAAAGCAAAAGAAAGACTACCAAAACCCAAACACAGAGAAAA	2196
Db	2244	GTTTGGGTTGAGAAAGAGAGTAAAGAAAGCCACAGATTAAGAAAATCTAAACCTAAAG	2303
Qy	2197	AGTATTAATTTTCAAGAAAGAGAAACCAATCCCTGCAATATGCTCTTGTTGTAATG	2256
Db	2304	AGACTTCA-----AAGCAGATTAATGATCGCTAGAGAAATGTCAAGAG 2345	

QY	2257	ACGAAAGTGCCTCCAGAGAGCTGAGAAATGAAAAGCCGGTATTGTAAATCAACAAAATTAAG	2316
Db	2346	GTTGTTATCGTCCCAAGTGTCAAATGTTGAGAAAGAGATCTGAAGCAACAAATTGAAATTGG	2405
QY	2317	AAAAGAAATTTGGGCAATCTTCTGTTTTTTTATCATCCACACTTCTCGAAGATGGTGGAA	2376
Db	2406	AGAAAGAAATTGGACAAATCTCCGGTTTTTCGTGCTCTGCTGTTCTACAGAACGGTGGAG	2465
QY	2377	CGTTGAAGAGTCAAGTCTGCTTCTCTTTTGAAGAAGCTATACATGTCATTAGTTGTG	2436
Db	2466	TTCCCCGGAACGCAAGCCCCGCAATGTTGTTAAGAGAAAGCAATCAAGTATTATGCTGCG	2525
QY	2437	GTTATGAAGACAAACAGACTGGGGGAAAAGAATGGCTGATCTATAGATCAATTAACAG	2496
Db	2526	GGTACCAAGATTAACCCGAATGGGGAAAAGAATCGGGTGAATTTATGATCGGTGACTG	2585
QY	2497	AAGATATTCTAACCTGGTTTTCAAGAATGCATTGTCAATGGTTGGCCGGTCAATTATCTGCATAC	2556
Db	2586	AAGATATCTTGACCGGGTTTTCAAGATGCATTTGCCATGATGAGATCTGTGTACTGTATGC	2645
QY	2557	CTAAACGGGTTGCATTCMAAGGTTCTGCACTCTGAAATCTTTAGATCGTCTTCCACAGG	2616
Db	2646	CTAAGCGTGAAGCTTTTAAAGAGATCTGCTCTATTAACCTGTGACATCGTCTTATCAAG	2705
QY	2617	TGCTTCGGTGGGCTCTTTGGGTCTATTTAGATCTTCTTCAGCAATCATTTGCCCTTTGGT	2676
Db	2706	TTCTACGTTGGGCTCTTGGGCTCTGTAGAGATTTTCTTGAGCAGACATTTGCCGATATGGT	2765
QY	2677	ATGGGATATGATGGGCGGTCTGAAATTTTGTGAAAATTTTCCATATCACTCCATCGGT	2736
Db	2766	ATGGTATATGGTGGGTTTAAATAGTTTGAAGAAATTTCTTATCATCAACTGTGTGCT	2825
QY	2737	ATCCTTGACACTATTTCCCTCTTGGCTTACTATCATTCGTCGACATCGTTATATGA	2796
Db	2826	ATCCTTGACCTTCACTTCCATTGATTCGTATCTGTCTCTCCCGCGTTTGTTTACTCA	2885
QY	2797	CAGGAAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTTATGTCACCTT	2856
Db	2886	CAGGAAATTTATCATGCTCCCTGAGATTAAGCACTACGCAAGATTAATCTTCAATGCTCATGT	2945
QY	2857	TTATCTGATTTTTTGTCTACAGAGCATCTCTAGAAATAGATGGAATGGATTTGAAATTAAG	2916
Db	2946	TCATATCATATGCAAGTAACTGGAAATCTCTGAAATGCAATGGGAGAGTTCGAAATCAATG	3005
QY	2917	ACTGGTGAAGAAATAGACAGTTCTTGGGCTATTGAGAGTGTCTCTCAACACTCTTGTGCTG	2976
Db	3006	ATTGGTGAAGAAAGACAGATTTTGGGTAAATCGAGAGGGGCTCTCTCGCATCTATTTGCTC	3055
QY	2977	TGTTCCAGGAGCTTCTCAAGGTCAATAGCTGTGTGTAACAAGCTTCAACCGTGACATCA	3036
Db	3066	TGTTTCAAGTTTGTGCTCAAGTTCTTAGCCGAGTTTAAACGAATTTCAAGTCACTTCA	3125
QY	3037	AGGTGAGAGATGATGAGAGATTTCTCAGAGCTATATCACTTCAATGCACTACCTTATGA	3096
Db	3126	AAGCAGCAGAGATGAGACTTCTCTGAGCTTTACATCTTCAAGTGAACAATTTGTGGA	3185
QY	3097	TACCTCTAACCAACCTGTCTCATATGAACCTCATTTGGTGTGATCGCTGAGCGTTTCAATG	3156
Db	3186	TTCTCTCCAGAACACTTCTGATCATTAACATCAATGAGATTATGTGCGCGTTTCTGATG	3245
QY	3157	CGATCAATPACGATATAGTCAATGAGGCCCCCTCTTGGAGAGCTATCTTGTGCAATTT	3216
Db	3246	CGATTATGCAATGGCATATGATCACTAGGGGACCTCTTTGGGAGACTTTTCTTCCGCTCTT	3305
QY	3217	GGGTGATTTGCTATTTTATTCCTTTCTCAAAAGTTTGGTTGGAAAGGCAAAACAGACAC	3276
Db	3306	GGGTATTTGCTATTTATACCATCTCTCAAGGAGATCTTGGGAGCAACAACAAATGTC	3365
QY	3277	CAAGCATTTGTAATGTCCTGTCGTCATTTGCTGAGCTTCAATCTTCCGCTCTTGGGTTC	3336
Db	3366	CTACAGATTAATGTGATCTGTGTCATTTCTTTACGTTCAATCTTGACACTCTTGTGGGTCA	3425
QY	3337	GGATTTGATCTTTCCTTCGAAAGATGATGTCGCTTTCTTGAGAGATGTGTTTGAAT	3396

Db 3426 GAATTAACCGGTTTGCTA---AAGGGGACCACTGTGGAGATCTGTGCTGAATT 3482  
QY 3397 G 3397  
Db 3483 G 3483

RESULT 12  
AF200526 3725 bp mRNA linear PLN 31-AUG-2000  
LOCUS Zea mays cellulose synthase-2 (Cesa-2) mRNA, complete cds.  
DEFINITION AF200526  
ACCESSION AF200526  
VERSION AF200526.1 GI:9622875  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoidae; Andropogonae; Zea.  
1 (bases 1 to 3725)  
Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S.,  
Xocoostle-Cazares, B. and Delmer, D.P.  
A comparative analysis of the plant cellulose synthase (Cesa) gene  
family  
Plant Physiol. 123 (4), 1313-1324 (2000)  
JOURNAL 20398328  
MEDLINE 10938350  
PUBMED 2 (bases 1 to 3725)  
REFERENCE Dhugga, K.S. and Helentjaris, T.G.  
AUTHORS Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred  
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
source Location/Qualifiers  
1..3725  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/note="Identified and isolated from the maize genome  
database cDNA collection of Pioneer Hi-Bred International,  
Inc"  
1..3725  
/gene="Cesa-2"  
179..3403  
/gene="Cesa-2"  
/note="Cesa-2"  
/codon\_start=1  
/product="cellulose synthase-2"  
/protein\_id="AA89962.1"  
/db\_xref="GI:9622875"  
/translation="MAKNGMAVAGSHNNEFVMI RHGDAPVPAPKPYKSGANGQVCOIC  
GDTVGSATGDFVACNCAFPVCRPCYEREGNCCPOCKTRKXKSGSPRVNG  
DEEDVDLDNENFKYKQNGKPEMQLQDDADLSSAHDPHRIPLTSGQISGE  
IPDAPDRSISPTSSYDPAVPEVRVDSKOLNSGLNSDVKRSHVSRVQD  
KMLQVNTKYPARGDMGEGSGENMOVDDARLPLSHVIPSQNLVYIIIL  
LIICFPFRYRISHPYRNAYGLMVSIVIERWPAISMLDQPKRYPIINRYLILRLA  
LHYRDEGSPQLAIDVSVYDPLKPEPLTANTVLSLQIDVLDYKDSVYSDGS  
AMUTFESLSTAFARKWVPCFKHNIIBRAPFEPYQKIDYKDKIQSPYKERAM  
KREYEFKIRIYNALVAKQKPEBGTMDGTAMPNNRDRHGMIOVLGSGGLDT  
DGNELRLVYSEKRGFOHKKAGANMALIRVSAVLNGAVLMDCHDFNSKA  
LRBAMFMDPALGRKTCVOPORFDGIDLDRVANRNIIVPDIIMKLDIOGVV  
VGTGCGFRNALYGNPVTLEADLEPNLYKSCGGRKKNKSYNDOSRIKRPSS  
APTFNEDIIEGIBGYEDERSVMSORLKRKRGSPITASTPMTQGGIPSTNPAS  
LKLEALHVISCGEYDEKTEWKEIGMTYGVTEDDILGFMHARGWQSYICMPRPCK  
GSAPINLSDRLNQVLMALGSVELLSRCPIMYGNGRLKLELAYINTIYVPTS  
VPLIACVLPALICLTKFKIPEISNACMFILFASIPAGIILADRYGVIIDBM  
RMOEFVIGTSAHLFAVFOGLKVLVAGIDTNTVTSKSDSEGDAELIYVFTSL  
IIPETVIVINUNGMAGISYAINSGVOSQSLFGKLPSTIWIILHPIPLKLMGRON  
KPTIVIVISILASTISFLMWKIDPFIPTQAAALGGCGVNC"

BASE COUNT 927 a 807 c 950 g 1041 t  
ORIGIN  
Query Match 35.0%; Score 1378.4; DB 8; Length 3725;

Best Local Similarity 66.4%; Pred. No. 3.6e-256;  
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;  
QY 119 CGCGGGAAGTGGAGGAGAGAAAGCATGAGCCGAGCCCGAGCTGTGTGCGGCTCCCA 178  
Db 154 CGGTGTGTAGAGAGAGCGGAGATGCGCCCAAGAGGAGATGGAGGCTCTCA 213  
QY 179 CAACCGCAACGAGCTGTGTATCATCCGCGGAGCGGCA-----TCCGGGCGAGAGCC 232  
Db 214 CAACCGCAACGAGCTGTGTATCATCCGCGGAGCGGCAACGCGCTGTGTCCGGCTTACGC 273  
QY 233 GCGCGGAGAGAGAAAGGAGAGAGTGGCAATTTGCGGCGAGAGAGTGGCTTGGCCCC 292  
Db 274 CACGAAGAGTGGGAATGGGAGAGTGTGCAGATTTGTGGCGACATGTTGGGTTTACGC 333  
QY 293 CGGCGGAGAGCCCTTCTGTGTGTGTGCAAGATGCGCTTCCCGTGTGTGCGGAGTGGCTA 352  
Db 334 CACTGTGTATGTCTTGT 393  
QY 353 CGAATACGAGCGCGGAG 412  
Db 394 TGAGTACGAGCGCAAG 453  
QY 413 CCTCAAGGCGTGGCAAGTGTGACCGGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472  
Db 454 ACAGAAAGGTAGCCCTCGAGTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 513  
QY 473 CAAGAGTTCATCTGGAGAGCGGCAATGACTCGCATCTGTGTGTGTGTGTGTGTGTGTGT 532  
Db 514 CATTAATTCATCACTTAA-----GCAAGCAATGAGAGAGG 549  
QY 533 CCAGATGAGCTAGCGGCGGTGAG 592  
Db 550 CCAGAGTGGAGAGCTTCAAG 609  
QY 593 CCCCAAT---GTTCCACTCTCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649  
Db 610 ACACATGAGAGATTCACAGCGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669  
QY 650 GCAGCGCGTGTGCTTCTTGT 709  
Db 670 ATCCCTGAGCGGTCTTCTATCCGAGT-----CCAACTGAGAGTA 711  
QY 710 TCGGATCCAGCTTACCTGTGCAACCCAGAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 769  
Db 712 TGTATCCAGAGCTTCAAGTCTGTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 771  
QY 770 ATATGGATATGATGT 829  
Db 772 CATAGGCTTATATGT 831  
QY 830 GAGATGACACAGAGGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 873  
Db 832 CAATAATATGT 891  
QY 874 -----ACAGT 925  
Db 892 TGGCTCAATGTGAGAGATGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 951  
QY 926 AATTCACCTTCATGACAGAGATTAATCATATGATGATGATGATGATGATGATGATGATGAT 985  
Db 952 TGTGCAATTTCTCAACAGAGCTCAACTTACCGGATAGTATGATGATGATGATGATGATGAT 1011  
QY 986 GGTTTTGGGTTCTTCTTCACTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1045  
Db 1012 CATCTGT 1071  
QY 1046 GTGGCTATGT 1105  
Db 1072 GTGGCTATGT 1131  
QY 1106 CCCAAGT 1165

D 1132 CCCAAAATGSTATCCATCAACCGTGAACATATCTCGACAGGCTTGCAATTGAGTATGA 1191  
Q 1166 CAGGAAGGCCAGCCATCTCAACTCTGCTCAATTGATTTCTTTGTCAGTACGGTTGATCC 1225  
D 1192 TAGAGAGGAGAGCCATCAACAGCTGGCTCCCATTTGATGTCTTTGTGACGTACAGTATCC 1251  
Q 1226 CTTAAAGGAACCTCTTGTGTGTACACAAATACGTGTCTATCTATCTTTGGGTGATTA 1285  
D 1252 ATTGAAGGAACCTCCACTGATCAACGCCAACACGTGTTTGTCCATCTTGTGTGATTA 1311  
Q 1286 TCCGTGTAAGAGTTCTTGCTATGTTTCTGATGATGTGTGCAATGCTAACGTTTGA 1345  
D 1312 CCTGTGACAAAGTGTACATGATGTTTCTGATGATGCTCAGCTATGCTTGACTTTTGA 1371  
Q 1346 AGCATTAATCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACGGTACA 1405  
D 1372 GTCTCTCTGAAATCTCCGAATTTGTCTAAGAAAGTGGTTCCTTTTGTAAAGACAA 1431  
Q 1406 TATTTGAACCTCGGCTCCAGAGTGTCTTCCAAAGAAATAGACTTACTTTGAAAGACA 1465  
D 1432 TATTTGAACCAAGAGCTCCAAATTTTACTTTGCTCAAAAAATGATTTACCTGAAGACA 1491  
Q 1466 GGTGCGCAAACTTTGTGTAAGAGAGAGCAATGAAGAGATGAGATGAGAAATTCAA 1525  
D 1492 AATTCAACTTCAATTTGTAAAGAAAGACAGCAATGAGAGATGAGAGATTCAA 1551  
Q 1526 GGTGAGATCAATGCTTAACTTGTGCAAGGCCAAGAAATTCCTGAAGAGATGACAT 1585  
D 1552 AATTGAATCAATGCTTGTGCTGCAAGGACAGAAAGTGTCTGAAGAGGGTGGACAT 1611  
Q 1586 GCAGATGGAACCCCTGCGCTGGAACAAATGTTCTGATCATCTGAGATGATTCAGGT 1645  
D 1612 GGTGATGGAACCTGCTGGCTGGGAATAACCTGAAGGACATCTGCGCATGATTCAGGT 1671  
Q 1646 CTTCCTTGGCAAGCGGAGGAGCTTGACTGTGAGGAAATGAATGCTGACAGATGGTTA 1705  
D 1672 GTTCTTGGGCAAGTGTGGGCTTGACTGACTGAGAAATGAATACCAAGTCTTGTCTA 1731  
Q 1706 TGTTCCTAGAGAAAGCAACAGGCTTAACATCATATGAAAGCTGTGCTATGATGC 1765  
D 1732 TGTCTCTGGAAGAAAGACAGGCTTTCAGCATCAAGAGGCTGTGCTCAATGATGC 1791  
Q 1766 ATTGGTCCGAGTCTCTGCTGTATCAAAATGCTCCATATTTGTAACTTGGATGTGA 1825  
D 1792 ACTGATTCGTGTATCTGCTGTGCAAAATGGCTATCTCTCATGTGATGTGA 1851  
Q 1826 TCACATCAACAACAGCAAGGCTATTAAGAAAGCAATGTTTATATGATGACCCCTT 1885  
D 1852 CCAATTCATTAAGCAAGAGCTCTTAAGAGCAATGTGCTTCAATGATGATCCAGC 1911  
Q 1886 ACTAGAAAGAAAGTTGCTATGATACAGTCCCTCAAAAGATTTGATGGATGATGCCA 1945  
D 1912 TCTAGGAAGAAACCTGTTATGTATCAATTTCCAAAGATTTGATGGATGATGCCA 1971  
Q 1946 TGACCCATATGCTAACCGGAATGTGCTTTTGTGATATCAACATGAAAGTTTGGATG 2005  
D 1972 CGATCGATATGCTAATGGAACATAGCTTTCTTTGATATCAACATGAAAGTCTAGATG 2031  
Q 2006 TATTCAGGCTCAATTAATGTTGTACTGATGTGATTTAAGAGGAGGATTAATG 2065  
D 2032 CATTCAGGCTCACTATATGTTGGAACAGATGCTGTTCAATAGGAGGCTTTGATGG 2091  
Q 2066 TTAATGATGCCCCCAAAACAAAGACCAATCAAGGACTTGCACCTGTGCGCCAAATG 2125  
D 2092 AATATGAT-----CCTGTTTGTAGAGCTATCTGGAACCTAACATTT 2134  
Q 2126 GTGCTTTGCTGTGCTGTGCTTGGCAATAGAGCAAAAGACATCAAAACCCAAAC 2185  
D 2135 GTTGTATAG-----AGCTGTGTGTAGAGAGAAAGAAACAAAGATTTATATGATG 2190  
Q 2186 AGAAGAAAGAAAGTATATTTTCAAGAAAGAGAACCAATCCCTGATATGCTCT 2245  
D 2191 TCAAGCCGTA-----TTATGAAGAGAACGAATCTTCAGCTCCATCTTTAACT 2241

Q 2246 TGTGAAATTTGACGAAGCTGCTCCAGAGCTGAGATGAAAAAGCCGGTATTTGTAATCA 2305  
D 2242 GGAAAGCATCGAGGAGGATTTGAAGTTATAGATGAAAGGTGAGTGTATATGCCA 2301  
Q 2306 ACAAATATGAAAAAATTTGGGCAATCTTCTGTTTTGTATCATCCACTTCTGA 2365  
D 2302 GAGGAATTTGAGAAACGTTTGTGATGCTTCAATCTTCAATGATCCACTTTATGAC 2361  
Q 2366 GAATGTGGAACCTTGAAGAGTCAAGTCTCTCTTTGAAGAAAGTATACATGT 2425  
D 2362 TCAAGTGGCATACCACTTCAACAAACCCACTTCTCACTGAAGAAAGTATCATGT 2421  
Q 2426 CATTAAGTGTGATTAAGACAAACAGACTGGGAAAGAGATTTGCTGATCTAGG 2485  
D 2422 TATCAGCTGTGGGTACGAGAACAAACCTGAATGGGAAAGAGATTTGCTGATCTATG 2481  
Q 2486 ATCAGTTACAGAAAGATTTCTAACTGTTTCAAGATGCTTTCATAGTTGGGCGCAT 2545  
D 2482 TTCAAGTTACAGAGATATTTCACTGGTTTAAATGCAATGCAAGAGCTGGCAATCAAT 2541  
Q 2546 TTAAGTATCCTAAACGGGTTGCATTCAAAGGTTCTGACCTCTGAATCTTTCAGATG 2605  
D 2542 CTAAGTCAATGCCAACACGACCTTGTTCAGAGGTTCTGACCAATCAATCTTCTGATG 2601  
Q 2606 TCTTACACAGTGTCTGGTGGCTCTTGGGTCTATATGATCTTCTTCAAGATCATTTG 2665  
D 2602 TCTTATACAGTGTCTCGTGGGCTCTTGGGTCAAGTGAATTCGTCTAGACAGCATTTG 2661  
Q 2666 CCTCTTGGTATAGGATAGTGGGCTGAAATTTTGGAAAGATTTTCTTACATCA 2725  
D 2662 TCTTATATGATAGTGTACAAATGGGCGATTTGAAGCTTTTGGAGAGCTGGCTTACATTA 2721  
Q 2726 CTCAATCGTATACCTTGGACATTAATCCCTCTTGTGCTATCTGATCAATGCTGCAAT 2785  
D 2722 CACATTTGTTATCCAAATACATCTGTTCCGCTTATCGCTATTTGTGTCTTCTGCTAT 2781  
Q 2786 CTGTTTATGACAGGAAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTT 2845  
D 2782 CTGTCTTCTTAACTAAATTAATATCAATCTGAGATTAATGATTAATGCTGAATGTTCT 2841  
Q 2846 CATTGACCTTTTATCTGATTTTGTCTACAGCAATCCAGAAATGAGATGAGATGCT 2905  
D 2842 CATCTCTTTTGTGCTCCATTTTGCACCTGATATTTGAGTCTAGATGAGATGCT 2901  
Q 2906 TGGAAATGATGACTGAGAGAAATGACAGTCTGGGCTCATTTGAGAGTGTCTCCACA 2965  
D 2902 TGGCAATGAAAGATTTGATGAGAAATGAGCAATTTTGGTATTTGGTGGACCTTCCCA 2961  
Q 2966 CTTCTTGTGCTGTCCAGGAACTTCAAGGTCAATAGCTGTGTTGATACAACTTCA 3025  
D 2962 TCTCTTGGGAGTGTCCAGAGGTCTGCTGAAGTGTGCTGGGATTTGATACCAACTTCA 3021  
Q 3026 CGTGAATCAAGG---GTGAGATGATGAGAGTCTCAGAGCTATATCAATTCAAATG 3082  
D 3022 AGTTACCTTAAGAGCATGATGAGATGAGCACTTTGTGAGCTATATGTTGTTCAAGTG 3081  
Q 3083 GACTACCTTATGATACCTCTTACCACTTGTCTTATTAAGAACTTCAATGATGATGCTG 3142  
D 3082 GACCAATTTGTCTATCCCTCCAGCACTGTTCTTGTCAATTAACCTGTGTGAAATGCTG 3141  
Q 3143 TGGCGTTCAAAATGCAATTAACGATATGATGATGAGGAGGAGGCTCTTTTGGAGGT 3202  
D 3142 AGGAATTTGATGCAATTAACAGCGGCTAACCAATCTGGGATTCGCTCTTTGGAAGCT 3201  
Q 3203 ATTCCTTGAATTTGGGATATGCTCATCTTATACCTTTTCAAAAGATTTGGTGGAGG 3262  
D 3202 GTTCTTCTGATCTGGGTATCTCCATCTATCACTTCTTCAAGGATCTCATAGGGAG 3261  
Q 3263 GCAAAACAGGACCAACGATTTGATCGTCTGTGCTCATTTGCTGGCTTCAATCTTCTC 3322  
D 3262 GCAGAACCGACCGCAACATGCTATCGTTTGTCTCATCTCTTGTGCTCTATCTTCTC 3321



OY	3323	GCTCCTTGGGATCGATTGATCCCTTCCT	3352
Db	3322	CTGTCTGTGGTGGAAGATCGATCCTTTCAAT	3351
RESULT 13			
LOCUS	AF200525	3752 bp	mRNA linear PLN 31-AUG-2000
DEFINITION	Zea mays cellulose synthase-1 (Cesa-1) mRNA, complete cds.		
ACCESSION	AF200525		
VERSION	AF200525.1	GI:9622873	
KEYWORDS	.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyte; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3752) Holland,N., Holland,D., Helentjaris,T., Dhuga,K.S., Xocoostle-Cazares,B. and Delmer,D.P.		
TITLE	A comparative analysis of the plant cellulose synthase (Cesa) gene family		
JOURNAL	Plant Physiol. 123 (4), 1313-1324 (2000)		
MEDLINE	20398328		
PUBMED	10938350		
REFERENCE	2 (bases 1 to 3752) Dhuga,K.S. and Helentjaris,T.G. Direct Submission Submitted (01-NOV-1999) Agrogenomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1..3752		
source	/organism="Zea mays" /mol_type="mRNA" /db_xref="taxon:4577" /note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"		
gene	1..3752 /gene="Cesa-1" 183..3410 /gene="Cesa-1" /note="Cesa-1" /codon_start=1 /product="cellulose synthase-1" /protein_id="AAF89961.1" /db_xref="GI:9622874" /translation="MAANKGVAVGSHNNEFMIRHDGDVBSAKPTKSANGVOICIGDSVSATGIDVFVAECNECAFPCVRCEYEKKEGNOCPOCKTRKYRKOKSPVHGSGDESDVDLDNEFMYKQSGPEMKGADLSSASRHPHRIPRLTSGOOISGEIIPDASPBRHSRSPSTSVSDSVVPPIIVPBSDINSYGINSIDMKERYSMRVKDDIKMMOVTNKTYEARAGDMEGESNGEKMQMTDARLPISRIVPSSNOLTRYVVITILRIILICEFPQVRSHPNRDAYGLMLVSTICEWPLSLWLDOQPKWPFINETYLDRALARDRESEBEPQLAPIDVFSTVDLPLEKRPILTANTVLSISDVYPDKSCVSDGOASAMTFDESIESTAEPARKMPVPCKKHNLNPAPAEFYFKQIDYKDIOFSEVEREARMKRYEEBEKVRIINALVAAKQVPEGMGMAAGTAMPNGPRDHGMTOVFLGHSGGLTDGNEELRVLYVSRERKRGQHKKKAGMANLIYSAUTLTGAYLANIDCHVYNSGSALRAEMCFMDPALGRKTCTVQFFORFGIDLHDIRARNIVFFDINKGLDGIOGPRIYVGGCFCRQALGYDVLTEADLEPIVIKSCGRRKRNTVMDSOSIMRTESSSAPLFNMEDIIEGLEGEDERSVLMOSQRKLKRCGOSPFIASFPTOGGIPSTINPESILSLKEAHLISCGVEDTKMEKEIGMIYGVSTEDILTFGKHARGMOSIYCMPPCPCKGSPINLSDRLNOLRWALSGVEILLERHSHOIVGVNGRLKLERLAYINTIYVPLETMSPIAACVCLPAICTLNKFLIIPRISNAWGPFLLIFATGILIELRWGVIDEOMWRNOCVIGCTSAHLPRVFOGLKTVLAGIDTNTYTYSKADEBGDPRAELYVPFTTSILIPFTVALINLVGWAGISTAINSGVSWPLFGKLFSTIVLHLTPFLKGLMGKROLIPRTVIVWSILLASTIFSLLMVXIDPISPTORAALAGQGVCNC"		
BASE COUNT	944 a	804 c	965 g 1038 t 1 others
ORIGIN			
Query Match	34.7%	Score 1364;	DB 8; Length 3752;
Best Local Similarity	66.1%	Pred. No. 2,2e-25;	
Matches 2165; Conservative	0;	Mismatches 1001;	Indels 111; Gaps 9;

OY	115	TCGGCGGGAAAGTGGAGAGGGAGAAAGCGATGAGAGGCGACGCGCGGCTGTGTGGCCGCT	174
Db	154	TCGGTGGAGGAGCGGGGAGGAGCGGAGATGGCGGCCAAGAGGAGTGTGTGGGGGCT	213
OY	175	CCCAACAAACCGCAACGAGCTCGTCGTCAATCCCGCCGCAAGCGCAT-----CCCGGGCCGA	228
Db	214	CGCAACACCGCAACGAGTTCTTCATGATCCGCCACGACGGCGATGTGCCGGGCTCGGCTA	273
OY	229	AGCCGCCGGGAGACGAAACGGGCAAGTGTGCCAGATTTGCGGCGACGACGTCCGACTTTG	288
Db	274	AGCCACACAAAGTGCAGATGGAAGAGCTCTCCAGATTGGCGGTGACTCTGTGGGTGT	333
OY	289	CCCCCGGGGGGACCCCTTCTCGTGGCGTGCACAGATGGCGCTTCCCGCTCTGCGGGAGCT	348
Db	334	CAGCCACTGGTGAATGTCTTTGTTCCTCGCAATGAGTGTGTCCTTCTCTGTCCGCCCAT	393
OY	349	GCTACGAATAGACGCGCGGGAGGGACACGACAACTGCGCCCAAGGCAAGACTGCATACA	408
Db	394	GCTATGAGTATGACCGCAAGAGAGGGAGCAATGTGTGCCCCAGTGCACAACTAATACA	453
OY	409	AGCCCTTCAAGAGGCTGCCAAGTGTGACCGGTGACGAGAGAGAGACGCGCTCGATTACC	468
Db	454	AGAGACAGAAAGTAGCCCTCGAGTTCATGGGTATGAGATGAGGAATATGTTGATGACC	513
OY	469	TGGACAAAGATTCATCTGGGACGGCCATGACTCGCACTCTGTGGCCGATTCATGCTCT	528
Db	514	TAGCAATGAAATTCACCTCAAA-----GCAGGCAATGGGGA	549
OY	529	ACGACCAATAGCTACGCGCGGTGAGGTGACCCCTAATGGGCGCGCAACAGTTTCC---	585
Db	550	AAGGCCCAAGATGGCACTGCAAGAGATGATGCTGATCTGTCTTCACTGTGTCGCCATG	609
OY	586	AGCTCAACCCCAATGTTCCACTCTCTCAACCAACGGGCAATGATGATGATCAATCCACGG	645
Db	610	AGCAACATCATCGGATTCACGCGTCACAAACGGGTCAACAGATATCTGAGAGATTTCTG	669
OY	646	AGCAGCACGCGCTGGTGCCTTCTTTCATGGGTGTGGGGAAAGAGATACATCCCTTTC	705
Db	670	ATGCTTCCCCCGAACCGTCAATTCATTCGCGCAT-----CCAACATCGA	711
OY	706	CTTATGGCGGATCCCGAGCTTACCTGTGCAACCCAGGTCTATGGAACCATCAAGATCTTG	765
Db	712	GCTATGTTGATTCAGAGCGTCCCAAGTTCTCTGAGGATGTGTGACCCCTCGAAGGACTTGA	771
OY	766	CTGCATATGGGTATGTAGTGTCTCTTGGAAAGAACGGATGGAGATTTGGAAGCAGAGAC	825
Db	772	ATTCCTATGGGCTTAATATGTTGTACTGAGGAAGAAAGTTGAGAGCTGAGCGCTTAAAC	831
OY	826	AAGAGAGATGCACACGACGGGGAATGATGCTGTGATGATGTG-----	873
Db	832	AGGACAAAAATATGATGCAAGTGACTAATAATATCCAGAGGCTGAGAGGACATGG	891
OY	874	-----ACGATGCTGATCTACCACTAATGAGATGGAAGCAACAACTGT	918
Db	892	AGGGGACTGCTCAAAATGGAANANATGCAAAATGTTGATGATGACAGGCTACCTTTGA	951
OY	919	CCAGAGAAATTCACCTTCATCAACGACGAGATTAATCCATATAGATGATTAATCATTTTC	978
Db	952	GCGGTATCGTCCAAATTTCTTCAACACGCTCAACCTTTACCGGTAAGTATCATTTCTC	1011
OY	979	GAGCTTGTGTTGGGTTCTTCTTCACTACGAGTATGATCATCCGCTGATGATGATCAT	1038
Db	1012	GTCCTTATCATCTGTGCTTCTTCTTCCAGATTCGATCAATCATCAATGCGTATGCTT	1072
OY	1039	TTGCTTTGTGGCTCATATCTGTATATCTGTAAATCTGTTTGGCATGTCTTGATTTCTTG	1098
Db	1072	ATGATATTAATGCTATGCTATGCTGATATCTGCAAGGCTGTGGTTTGCCTTGTCTGGCTTACG	1133
OY	1099	ATCAATTTCCCAATGTGTTCCCTATTAGAGAGAGACTTACCTGACCGGCTGTCACTGA	1156
Db	1132	ATCAGTTTCCCAAAATGATATTCATCAACGTGAGACATATCTTGACAGGCTTGATTTGA	1199
OY	1159	GGTTCGACAAAGGACCAACCATCTGCACTTGTCTCCAAATGATTTCTTGTGTCAGTACGG	1218

Db 1192 GGATATATAGAGGGAGAGCCATCAAGCTGGCTCCATGATGCTTCGTCAGTACAG 1251  
Qy 1219 TTGATCCCTTAAAGAACTCTTTGGTCAACAATACTGTTCTATCTATCTTCGG 1278  
Db 1252 TGGATTCATTGAAAGAACTCCCATGATCAGCCCAACACTGTTTGTCCATCTTCG 1311  
Qy 1279 TGGATATCTCTGTTAAAGTTCTTGCTATGTTTCTGATGATGCTGCATGCTAA 1338  
Db 1312 TGGATTAACCTCGTGAAGAGTGCATGCTATGTTCTGATGATGCTTCAAGCTATCTGA 1371  
Qy 1339 CGTTGAAGCATATCTGAAACATCTGAATTTGGCAAGAAATGGGTCCTTCTGAAAC 1398  
Db 1372 CTTTGAAGTCTCTAGAAACCGCAGAAATTTGCTGAAGTGGGTTCCCTTTTGAAG 1431  
Qy 1399 GGATCAATATTTGAACCTCGGCTCCAGAGTGTACTTCCAAACAGAAATAGACTAGT 1458  
Db 1432 AGCACAATTTGAACCAAGCTCCAGAAATTTTACTTTGCTCAAAAATAGATTAAGCTGA 1491  
Qy 1459 AAGACAAGGTGCAGCAAACTTTGTTAGGAGAGAGCAATGAAGAAGATAGAG 1518  
Db 1492 AGACAAAAATTCACCTTCATTTGTTAAGGAAAGACGCGCAATGAAGGAGATAGAA 1551  
Qy 1519 AATTCAAGTGAATCAATGCTTAGTTGCCAAAGCCAGAAAGTTCCGTAAGAAAGAT 1578  
Db 1552 AATTCAAGTGAATCAATGCTTAGTTGCCAAAGCCAGAAAGTCCGTAAGAAAGAGT 1611  
Qy 1579 GGACAATGCAAGATGGAACCCCTGGCTGGAACCAATGTTGCTGATCAATCTGAAATGA 1638  
Db 1612 GGCATATGCTGATGGAATGCAATGCTGGGAAATATCCTAGGACCAATCTCGGATGA 1671  
Qy 1639 TTCAGGTTCTTCTTGGCCAAAGCGAGGCTTACTGTGAGGGAATGAATGCAATG 1698  
Db 1672 TTCAGGTTCTTCTTGGGCAAGAGTGGGCTCGACACTGATGGAATGAGTTACCACTGC 1731  
Qy 1699 TGGTTATGTTTCTAGAGAAACGACCGAGGCTATACATGCTAAGAAAGCTGGTGA 1758  
Db 1732 TTGTCATGCTCTCGTGAAGAAAGACCGAGCTTTCAGCATCAAGAAAGCTGGTGA 1791  
Qy 1759 TGAATCATTTGCTCCAGTCTCTGCTGTAACAATGCTCCATATTTGTTAACTTGG 1818  
Db 1792 TGAATGCGTGAATTCGTGATCTGCTGTGTCACAAATGCTGCTATCTTTCACATG 1851  
Qy 1819 ATTTGATCATCTACATCAACACAGCAGGCTATAAGAGCAATGCTTTTATGATG 1878  
Db 1852 ATTGCAACCATTACTTCAATAGCAGCAAGCTCTAGAGCAAGCAATGCTTCATGATG 1911  
Qy 1879 ACCCTTACTAGAAAGAGTTTGTATGTACAGTTCCCTCAAAATTTGATGAGATG 1938  
Db 1912 ATCCGGCTTAGAGAAAGAAACCTGTTATGTACAAATTTCCACAGAAATTTGATG 1971  
Qy 1939 ATGCGCATGACGATATGCTAACCGGAATGTTGCTTTTGTATCAACATGAAGTT 1998  
Db 1972 ACTTGAACGATGATATGCTAATCGGAACATAGTTTCTTTGATATCAACATGAAG 2031  
Qy 1999 TGGATGATATTCAGGCTCAATTTATGTTGTAATGATGCTGATTTAAGAGCAGCAT 2058  
Db 2032 TGGATGAGCATTCAGGCTCCAGTTTACGTGGAAACAGATGCTTCAATAGCAGGCT 2091  
Qy 2059 TATATGTTATGATGCCCCCAAAACAAAGACCATCAAGACCTTGCACTGCTGGC 2118  
Db 2092 TGTATGATATGAT-----CCTGTTTGACTGAAGCTGATCTGAGGCT 2134  
Qy 2119 CCAAGTGTGCTTTTCTGTTGCTGCTTGGCAATAGAAACAAAGAAAGACTACCAAC 2178  
Db 2135 AAACATTTGTTATTAAG---AGCTGCTGTGTTAGAAAGAAAGAAAGAAAGAAAGTTA 2190  
Qy 2179 CCAAAACAGAGAAAGAAAGTATATTTTCAAGAAAGAGAAACCAATCCCTGCAT 2238  
Db 2191 TGGATAGTCAAAACCGTA-----TTATGAAGAGAACAGATCTTCAAGCTCCCATCT 2241  
Qy 2239 ATGCTCTGTGTAATTTGACGAAGCTGCTCCAGACCTGAGATGAAGAGCCGATTTG 2298

Db 2242 TCATATGGAACATCCAGAGGCTATTGAAGTTACAGAGATGAAGATCAGTCTTA 2301  
Qy 2299 TAAATCAACAAAATTTAGAAAAAGAAATTTGGCCAACTCTGTTTGTATCATCCACAC 2358  
Db 2302 TGTCCAGAGAAATTTGAGAAAGCTTTGGTCAAGTCTCTATTTTCAATTCATGATCACT 2361  
Qy 2359 TTCTGAGAAATGATGGAACCTTGAAGATGCAAGTCTGCTTCTTTTGAAGAAGCTA 2418  
Db 2362 TTATACACAAAGTGGCATACCACTTCACAAACCCAGCTTCTTCAAGAAAGATCTA 2421  
Qy 2419 TACATGTCATTAGTTGAGTTATGAAACAAAGACAGACTGGGAAAGAAAGATGGCTGA 2478  
Db 2422 TCCATGTCATCAGTTGTGATGATGAGCAAAACCTGAATGGGAAAGAAAGATGGCTGA 2481  
Qy 2479 TCTATGATCAAGTTACAGAAATATCTPACTGTTTCAAGATGATGCTGATGTTGGC 2538  
Db 2482 TCTATGTTCAAGTACCGAGGATATCTGACTGGGTTTAAATGATCCAGAGGCTGGC 2541  
Qy 2539 GGTCAATTTACTGCATACCTAAACGGGTTGCATTTAAAGTTCTGACCTGAACTCTT 2598  
Db 2542 AATCAATCTACTGCATGCAACGACCACTTGTTCAGAGGTTCTGCAACCAATCAATCTT 2601  
Qy 2539 CAGATCGTCTTCCACAGGCTGCTGGGCTGCTGGGCTATGAGATCTTCTTCAACA 2658  
Db 2602 CCGATCGTCTTATATAGAGTGCCTCGGTGGGCTTTGGGTCAGTGAATTTCTGTTAGTA 2661  
Qy 2659 ATCATTTGCCCTTTTGTATGAGGATAGGTGGCGGTCTGAAATTTTGGAAAGATTTTCT 2718  
Db 2662 GACATTTGCTATCTGATATGATGTTCAATGACATGAGCTTTGAGAGGCTGGCTT 2721  
Qy 2719 ACATCAATCCATGCTGATATCTTGAACATATTCCTCTTGGCTTACTGTACATTTGC 2778  
Db 2722 ACATCAACATATTTATATCCAAATCAGATCATTCCTGCTATTTGCTGCTTC 2781  
Qy 2779 CTGCCATCGTTTATTTGAACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCA 2838  
Db 2782 CCGCTATCTGCTCTTACCAATTAATTTATCAATCTGAGATTTAGCAATTAATGCTGGA 2841  
Qy 2839 TGTGTTCAATGTCATTTTATCTGCAATTTTGTCTACAGACATCTTCAAGAAATGAGATGA 2898  
Db 2842 TGTTCCTCATTTCTTTTCCGCTCCATTTTGGCCACATGATATATGAGGCTTATATGGA 2901  
Qy 2899 GTGTGTTGGAATTTATGATGCTGTGAGAAATGACAGTTCTGGGTCAATGGAGTGT 2958  
Db 2902 GTGGTGTGGCAATTAAGATTTGTGGAAGAAATGACAGTTTGGGTTATTTGTGCACT 2961  
Qy 2959 CCTCACACCTCTTGTCTGTTCCAGGACCTTCTCAAGGTCATGCTGTGTTGATACA 3018  
Db 2962 CTGCCCATCTCTTCCAGATGTTCCAGGCTGCTGTAAGTGTGGCTGGATGATACCA 3021  
Qy 3019 GCTTCAACGTCATCAAGG---GTGAGATGATGAGAGTTCTCAAGCTATATACAT 3075  
Db 3022 ACTTCAAGTTATCTCAAGGATCTGATGAGATGGGAGCTTTGCTAGCATATGATG 3081  
Qy 3076 TCAATGATCTACCTTATTTGATCTCTTACCACTTCTGCTATATGAACCTTCAATGCTG 3135  
Db 3082 TCMACTGACCAAGTTGCTCATCTCTCCGACACAGTCTTGTCTATTAACCTGCTGGA 3141  
Qy 3136 TGGCTGCTGGGCTTCAATGATCAATTAACGATATGAGATGAGGAGGAGGAGGAGGAGG 3195  
Db 3142 TGGTGGAGAAATTTCTTATGTCATTAACAGTGGCTTCAACATCTCGGAGTCCGCTTTG 3201  
Qy 3196 GGAAGCTATCTTTCGATTTTGGGATGTCATCTTATCCCTTCTCAAGAGGTTTGG 3255  
Db 3202 GAAAGCTGTTCTTTCGATCTGGGATGTCCTCACTCTCAACCTTCTCAAGAGGTTCTCA 3261  
Qy 3256 TTGGAAGCAAAACAGACAGCAACAGATTTGATGCTGTGCTCATTTCTGCTGGCTTCAA 3315  
Db 3262 TGGGAAGGAGAAACGACACCAACATGTCATTTGTGTCATCTTCTTGTGATCTA 3321  
Qy 3316 TCTTCTGCTCTTCTTGGGTTGAGTTGATCTTCTCT 3352  
Db 3322 TCTTCTCTGCTGTGGGTGAAGATGATCTTCTCAT 3358



```

/translacion="MEPHLHARRRRKNFTIRLQDDQLAVSHDELQELASSHFGL
GISVERFDALDLHSLG/CPAHLGLDNPSESEVMAVVKALPDGKAPGDGTGRYK
TWSVVKPEVMAALNSLFLANGFSSLNDLILSLPKDEADVDFRPISPHSFG
KLFSKIPASRLSPHLGELVAPNQSASKADPFHNDPRYVQLAARALAHKIPILLIK
DLKAFDTGDLPLSPMLFLVMDVLTMTLRKASGGGLPLNDHASLYADLVFLSPS
VLSIPVHVSMTGLPAMMIKAIKKRRKRAFLMTGNVYMGELKAPARATDMQASTP
FVVDSESLTFMSDQMTIGRSVASLAPDLLFVPPRLRSRTVSSGLANNAMISDIRG
ALTVPGISQELV"
/gene="OSJNBA0042109.4"
/note="EST BF475198, BF475196, BE039861 from this gene"
/join("<12826, 13909, 13348, 13725, 14212, 14361, 14548, 14627,
14713, 14866, 15074, 15165, 15211, 15260, 15347, 15586)
/gene="OSJNBA0042109.4"
/join("12826, 12909, 13348, 13725, 14212, 14361, 14548, 14627,
14713, 14866, 15074, 15165, 15211, 15260, 15347, 15363)
/gene="OSJNBA0042109.4"
/codon_start=1
/product="unknown protein"
/protein_id="AA041135.1"
/db_xref="GI:28376705"
/translacion="MOLNPKARTRRNKTQKRVAMRRDCIERDLVWRMTSDQRYSAH
SAVOAFPLGHSFACVDLLANAKPAKCFELMFARQRCMMADLLQKSIDSHSTCP
FCOAELETAHLLIDVFAQVWLRYLSPGMNALSPPVAGSRTGHLPLLAAPVAG
EVPKRIAGIKALEFYLHNSITOSVHGDVKSNNILSEDPQQLCDPLGAKOVASST
PRLCTDITGTFGLAFEYVSHGVNKKIDVAFGVILVLEIISGRHRTICAKQ
ESLVGNAPKPLSSGEIKQLVDPFLGNDYDCDEMRMILAMHKDI FSPFGNVTDFETA
"
repeat_region 16877..16942
/rpt_family="tandem"
repeat_region 16879..16944
/rpt_family="CT-rich"
repeat_region 17388..17570
/rpt_family="tandem"
repeat_region complement(18185..18241)
/rpt_family="Oryza"
repeat_region 18186..18246
/rpt_family="Oryza"
repeat_region 18186..18243
/rpt_family="Oryza"
repeat_region 18186..18241
/rpt_family="Oryza"
/misc_feature 18186..18241
/note="Region with dinucleotide (AT) repeat (28-29
copies)"
repeat_region 18186..18241
/rpt_family="tandem"
repeat_region 18332..18414
/rpt_family="tandem"
repeat_region 18340..18412
/rpt_family="Oryza"
repeat_region 18343..18414
/rpt_family="Oryza"
repeat_region complement(18346..18414)
/rpt_family="Oryza"
repeat_region 18348..18420
/rpt_family="Oryza"
repeat_region complement(18348..18415)
/rpt_family="Oryza"
/misc_feature 18348..18413

Query Match 33.9%; Score 1335.4; DB 8; Length 171376;
Best Local Similarity 67.58; Pred. No. 1e-247;
Matches 2575; Conservative 0; Mismatches 456; Indels 786; Gaps 18;

QY 893 AATGATGAAGCAAGACAACTGTCAGGAAATTCACATTCATCAAGCCAGATTAA 952
DB 139935 AATGATGAAGCTGACAGCATGTCTCAGAAAGATTCATTCCTTCAAGCCGCGCAA 139876
QY 953 TCATATAGATGATATATCATTAATTCGGCTTGGCTTTGGGGTTCTTCTTCACTACCG 1012
DB 139875 TCCTTATAGATGATATATCATTAATTCGGCTTGGCTTCTTCTTCACTACCG 139816
QY 1013 AGTATGATCCGCTGATGATGATTTGCTTGTGCTCATATCTTTATCTGTGAAT 1072

```

```

DB 139815 AGTATGATCCAGTCCGCTGATGATTCATTTGATTAAGGTCATATCCGTCATCTGTGAAT 139756
QY 1073 CTGGTTTGGCATGCTCTTGATATTCCTTGATCATATTCCTTCAAGGTTCCCTATTTGAGAGAA 1132
DB 139755 CTGGTTTGGCATGCTCTTGATATTCCTTGATCATATTCCTTCAAGGTTCCCTATTTGAGAGAA 139696
QY 1133 GACTTACCTAGACCGGCTGTCACTG----- 1157
DB 139695 AACTTACCTTGAACCGGTTGACATGAGGTTGTACTCTTGGCACCTTGTATCACTTGGCC 139636
QY 1158 -----AGTTTCGACAAAG 1170
DB 139635 TTCTATTTCTTCCGAGGTTCTGTGACTTACTATTTTCTTCCGATAGGTTTGACAAAG 139576
QY 1171 AAGCGCAGCATCTCAATCTGTCATTTGTTCTGATGATGAGTGCATGCTTGAAGCAT 1230
DB 139575 AAGCGCAGCATCTCAATCTGTCATTTGTTTGTGATGAGTGCATGCTTGAAGCAT 139516
QY 1231 AGGAACCTCTTGGTGCACAAACAATACTGTTCTATCTTCTTCCGTTGATTTATCTCTG 1290
DB 139515 AGGAGCTCTCTTGGTGCACAAACAATACTGTTCTATCTTCTTCCGTTGATTTATCTCTG 139456
QY 1291 TTGATTAAGGTTCTTGGTGCATTTGTTCTGATGATGAGTGCATGCTTGAAGCAT 1350
DB 139455 TCGACAAAGTCTCTTGGTGCATTTGTTCTGATGATGAGTGCATGCTTGAAGCAT 139396
QY 1351 TATCTGAAAATCTGTAATTTGCAAGAAATGGGTTCTTCTGCAACCGTACATATTTG 1410
DB 139395 TGTCTGAAAATCTGTAATTTGCAAGAAATGGGTTCTTCTGCAACCGTACATATTTG 139336
QY 1411 AACCTGCGCTCCAGAGTGTGTTCTTCCACAGAAATGATCTATCTTGAAGACAGTGTG 1470
DB 139335 AGCTTGTGCTCCAGAGTGTGTTCTTCCACAGAAATGATCTATCTTGAAGACAGTGTG 139276
QY 1471 CAGCAAACTTTGTTAGGAGAGAGAGCAATGA----- 1503
DB 139275 CAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGTACATTGCAATTTCTGTGTGTTT 139216
QY 1504 -----AGA 1506
DB 139215 GCTATGATATGTAGTCATATAATTAATGCTAATTTTGAACCTCAATTCCTTCTCTCAGA 139156
QY 1507 GAGAGTATGAGAAATTCAGAGTGAATTAATGCTTACTTATGCGCAAGCCGCAAGATTC 1566
DB 139155 GGGAAATATGAGAAATTCAGAGTGAATTAATGCTTACTTATGCGCAAGCCGCAAGATTC 139096
QY 1567 CTGAAGAAGATGACAAATGCAAGAAAGGAAACCCCTGAGTGAAGCAATGTTGCGTATC 1626
DB 139095 CTGAAGAAGATGACAAATGCAAGAAAGGAAACCCCTGAGTGAAGCAATGTTGCGTATC 139036
QY 1627 ATCTTGAAATGATTC----- 1641
DB 139035 ATCTTGAAATGATTCAGTATGCTTTTACAGATTCATGATGCTGTTAATTTATCTGCTGAC 138976
QY 1642 ----- 1641
DB 138975 ATGATGATGAGTACTTCCCTTCTCTTGTCTTACATGACAAATCTCAACTTTTGC 138916
QY 1642 -----AGGCTTCTCTTGGCAAGAGGAGGCTTGAAGTGAAGGAAATGAAGCTGCAAC 1695
DB 138915 AACCATATAGTCTTCTCTTGGCAAGAGTGTGCGCATATGTGAAGAAATGAAGCTGCTC 138856
QY 1696 GATTGTTTATGTTTCTTGAAGAAAGCAAGGCTATTAACATCTAATAAGAAAGCTGTGTG 1755
DB 138855 GATTGTTTATGTTTCTTGAAGAAAGCAAGGCTATTAACATCTAATAAGAAAGCTGTGTG 138796
QY 1756 CTATGATGATTC----- 1768
DB 138795 CTATGATGATTCGTTAAAGACAAACGTTCAAGGATTTGTCAGTCTTCTATAGCTGTG 138736
QY 1769 -----GATCGAGATCTC 1780
DB 138735 AGCAGATCAGTGAATTTGATGATGCTTCTTGTGTCTGTTTTCAGGTCGAGTCTC 138676

```



Db	136517	CTGCATTTTAGCAACCGGAGGCTCGACA- CCTGTGAGTTCCGCCACGAAAAAT	136459
QY	3607	ACAGCATATGCAAGTGTGATTTGTCAGACATTTTATTAATCTGTGCGAATATAGATG	3666
Db	136458	ACAGTGAATGCCAGTGTGATTTGATTCACACATCTTTATTTAGTTGGTGGAAATTTGATC	136399
QY	3667	GGCTGAGCCGAACAGCAAGATATTTTGAATTCGACTGCTCCCGTGTACAAC-----	3719
Db	136398	GGGTGAGCCGAACAG-AAAAGTGTGATTTGTCATGTCTGTCCACATCTCAAAAATCT	136340
QY	3720	--TTGGTCTCAATTAAGCAGGCAAGATATGCAATCTGCC-AGTGGAAACAAGCAACTGCA	3776
Db	136339	CGGTACTGTCTTCATTAAGCAGCAAGAAAGATCTGCCCTCTCGAATTAGAGGAAGATGCA	136280
QY	3777	CATATATTATGA-----TGCCTGTCAATGAGGCGCTGTTCATTAATGATGCGTCGA	3830
Db	136279	CATTATTTATGTAGCGAGTTAAATATATGCTTAAGGCTGTTCATTTCTGTCTCTCTT	136220
QY	3831	TACTAGAAAAACAGATATTTAGCATTTATCTATGTTAATTAATTAAGATATGTAATGCGCC	3890
Db	136219	TCTATTTCAATTTTTTTTGTCTCTGTAGAACAAAATGATCTCTAGATTTAGTGAATTAAGCC	136160
QY	3891	TGTTTTTGTGTGACTGTATCAATCTGATGTGTT 3927	
Db	136159	CATCCGCTGTCTAATTTGTATAAGGTGCCCTTTGTT 136123	

RESULT 15				
BT008654				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
BT008654	3763 bp	mRNA	linear	PLN 16-MAY-2003
Arabidopsis thaliana clone RAFL09-gg-G08 (R19778) putative cellulose synthase catalytic subunit (RSM1) (At4g32410)		mrna,		
complete cds.				
BT008654				
BT008654..1	GI:30794042			
FLI CDNA.				
Arabidopsis thaliana (chale cress)				
Arabidopsis thaliana				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.				
1 (bases 1 to 3763)				
Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carrinci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
Arabidopsis Full Length cDNA Clones				
Unpublished				
2 (bases 1 to 3763)				
Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carrinci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
Direct Submission				
Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carrinci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				
The Salik, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.				

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers

```

1. 3763
   /organism="Arabidopsis thaliana"
   /mol_type="mRNA"
   /db_xref="taxon:3702"
   /chromosome="4"
   /clone="RAF109-89-G08 (R19778)"
   /note="This clone is in a modified pBluescript vector (pUC-1) as a BamHI/XhoI insert.
ecotype: Columbia"
1. 3763
   /gene="Atg32410"
1. 144
   /gene="Atg32410"
misc_difference
2
   /gene="Atg32410"
   /note="compared to genomic sequence"
   /replace="t"
145. 3390
   /gene="Atg32410"
   /codon_start=1
   /evidence=experimental
   /product="putative cellulose synthase catalytic subunit (RSM)"
   /protein_id="AAP0467.1"
   /db_xref="GI:30794043"
   /translation="MEASAGLVAGSVARNLTVRIRHSDGTVPLKMNNGOICGSE
DVLAEKGVFVACNBEAPVRCPCRYTEKDGTCQCCPKTRFRRRGSRVSGDEB
EDVVDIENFNTAQQAGNKAHQRHPRGFEVSSSHESQPIPLTHGTVGEITPDT
QSVRTTSGPSPDRNAISSPYIDPRQPVFVTVDPKDNLSYGLGVMDKERVGMK
LKEKMLQMTGKHEKGGEIGTSGNGBELQADDTLRPMSRVPIPSRLPYRRK
VILRLILCFPLQVTRTHPVKNAYPLMLNSVCEIWFAPSMISDQPKVPIRETRY
LBDLIRYDDRGESOLVPYDVFSVTDPLKEPLVYANVLSLSDYDPDKACV
SDGSAMLTRESLSETHPEAKKVPFEKKNTIEBRAPETPAQKIDYLDKDTFSV
ERAMKREYEEFVRINALVAKAQKIPBEEMTQDGPWPGNNTRDHPGMIQVLEGHS
GGLDQDNELPRLIYVSREKPGFOHKKKAGANALIRVSAVLNGLVILNDCDHYF
NSKAKIKEMAFMMDPAIGKCCCVOPQPGFDGLDHRVANNRIVPFDIMKGLDGI
QGVVYTGCCCFNRQALYGYDPLTEEDLEPNIIVKSCGSKKSKSKKNYKRRG
INSDSNAPLNNMEDIDEGREYDDESLMSQSVSKRGSPVPIAATMEGGELP
PTNPATLLKEAITHVISCYEDKTEWEKELGWTGISTEDILTGFKNARWISTYCK
PPRPAFGSAPINLSDRLNQLVRLMALGISIILSRHCPWYGHGRLLERLAVIN
IVPISIPISPIAACILPAPCLITDRFLIPEISNVASIMFILFISAVTGILBESWG
VSLIEDWRNQQFWITGGTSAHLFAVFGGLVLAGIDITNFTVSKANDDEDFAELVY
FKMTALIPPTVLNVLIGIVAGVSAVNSGWSQPLFGKLPALWVIAHLVPLFK
GLGRNRPTIIVMSVLAISIFSLMWRINPVDANPNANNNGKGVF"
3391. 3763
   /gene="Atg32410"
misc_difference
3747
   /gene="Atg32410"
   /note="not present in genomic sequence"
BASE COUNT      997 a      768 c      892 g     1106 t
ORIGIN
Query Match      33.9%; Score 1332.4; DB 8; Length 3763;
Best Local Similarity 65.0%; Pred. No. 2.8e-247;
Matches 2117, Conservative 0; Mismatches 1066; Indels 75; Gaps 7;
127 GTGAGGGGAGGAGCAATGAGGCGGAGCGCCCGGCTGTGCGCGGCTTCCACAACGCA 186
    |||
128 GTGTGTGGGTGGCTGCATGAGGCGAGTCCCGGCTTGTGCTGTGATCTCCACGAGAA 187
    |||
187 ACCAGCTCGTCGTATCCGCGGAGCGGCATCCCGGAGCGGAGCCGCGGAGGAGAGA 246
    |||
188 ACAGAGCTCGTTCGATCCGACATGAACTGTATGGCGGAGACCAACCTTTGAAGATATGA 247
    |||
247 ACGGGAGGTGTGCGGCGGAGCAAGCTGTGGCCTTGGCCCCCGGCGGAGACCTT 306

```

D	248	ATGGCCGATATATGTCAGATCTGTGTGATGATGATTGGACTGCTGGAACCTGGAAATGCTT	307
Q	307	TCGTGGCGTGCACACGAGTGCACGCTTCCCGCTCTGCCGGGAGCTGCTACGAATACGACGCC	366
D	308	TTGTCCGCGTGAATGAATATGTGCTTCCCTGTGTGTCCGCTTGTGATGATGATACGAGAGA	367
Q	367	GGGAGGGCAGCGCAACCTGCCCCCAGTGCACAGATCTGATACAGCGCTCAAGGCTGCC	426
D	368	AAGATGGAACTCAGTGTGGCCCTCATCAAGACTAGATTGACAGACACAGGGGAGTCT	427
Q	427	AACGTTGACCGGTGACGAGGAGGAGGACGGGTGTGATGACCTGGACACAGAGTTCACT	486
D	428	CTCGTTGTTGAAGAGATGAATAGTGGATTGATGTTGATGATATCGAAGATGAGTTCA---	484
Q	487	GGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTGACGGCCACATGAGTACG	546
D	485	-----ATTAGCCCAAGGGAGCTAACAGGCGAGACACCAAGCCA---TGGCAAG	532
Q	547	GCCGTGAGGTGACCTCTAATGGCCGCCACAGACTTTCAGTCAACCCCAATGTTCCAC	606
D	533	AGTTTCTTCTCTCTCTGACATGATATCTCAACCAATCTCTTCTCAACCATGGCCATA	592
Q	607	TCCTCACCAAGGGGCAATGTGTGATGATATCCCAACGGAGACGACGGCGGTGCTT	666
D	593	CGGTTTCTGAGAGATTGCAACGCTGTATACCAATCTGTGGAACTACATCAGGTCTT	652
Q	667	CTTTCATGGGTGTGGGGGAAAGAGATATCCCTCTCTTATGCGATCCAGCTTAC	726
D	653	-----TGGTCTCTTCTGACAGGAATGCTATTTCATCTCCATATATGATCCAGCGAAC	706
Q	727	CTGTGCACCAACGAGTCTATGGAACCCATCCAGAGATCTTGTGCATATGGATGTAGT	786
D	707	CTGTCCCTGTAAAGATCGTGCACCCGTCAAAAAGCTTGAACCTTATGGGCTGGTATATG	766
Q	787	TTGCTTGAAGAAGAACGATGGAATTTGGAAACAGAGACAGAGAGAGATG-----	836
D	767	TTGACTGGAAAGAAAGATGTAAAGCTGGAAGCTGGAACAGAGAGAAATATGTTACAGA	826
Q	837	-----CACCAAGCGGGGAATGATGGTGTGTGATATGATGGTACGATG	879
D	827	TGACTGGTAAATACCATGAAGGAAAGAGAGAAATTTGAAGGACCTGTTCCAAATGGCG	886
Q	880	CTGATCTTACCACTAATGGATGAGACAGACAACTGTCCAGGAAATTCACATTCAT	939
D	887	AAGAATCTCAATGGCTGATGATACAGCTCTTCATATGATGTGTGTGCTATCCAT	946
Q	940	CAAGCCAAATTATCCATATAGATGATTAATCATTAATTCGGCTTGTGGGTCT	999
D	947	CTTCTCGCTAACCCCTTATCGGGTGTGATTAATTCGCGCTTATCATCTGTGTCT	1006
Q	1000	TCTTCCATACCGAGTGTGATCCGCGTAATGATGCAATTTGCTTGGCTCATATCTG	1059
D	1007	TCTTGCAATATCTGACACTACCTGTGAAAAATGCAATCTTTGTGGTTGACCTGG	1066
Q	1060	TTATCTGTGAATCTGTTTGCATATGTCCTTGGATTCTTGATCAATTTCCAAAGTGTCTC	1119
D	1067	TTATCTGTGAATCTGTTTGCATTTTCTTGGCTTCTTGATCAGTTTCCAAATGATAC	1126
Q	1120	CTAATTGAGAGAGACTTAACCTTACCGGCTGTCACTGAGTTTGACAAGAGGCCAGC	1179
D	1127	CCATTAAACAGGAGACTTATCTTGACCGCTCTGCTAATAAGATATGATGACACGGTGAAC	1186
Q	1180	CATCTCAACTGTCTCAATGATATTTCTTTGACAGTACGGTGTATGCCCTTAAGAACCTC	1239
D	1187	CATCACACCTGTTCTCTGTATATGTTGTGTATGACAGTGGACCATTTGAAAGGCTTC	1246
Q	1240	CTTGTGTACAACTACTGTCTCATCTTATCTTCCGTGGATTAATCTTGTGATAAG	1299
D	1247	CCCTTGTATACGCAACACAGTCTCTGATTTCTTGTGTGACATACCCGGTAAATTAAG	1306
Q	1300	TTTCTGTATATGTTCTGATATGAGTGTGCAATCTAACGTTTGAACATTAATCTGAA	1358

Db	1307	AAGCCTGTAATGTTTCAGATGATGAGTTCAGCTATGCTTACCTTTGGAAATCCCTTCTGAAA	1306
OY	1360	CATCTGAATTTGCAAGAAATGSGTCTCTTTCTGCAAAACGGTACATATTTGAACTTCGCG	1419
Db	1367	CCGCTGAGTTTGGCAAAAGAAATGGGTGCACATTTTGSCAAGAAATTCAAATTTGAAACCTTAGGG	1426
OY	1420	CTCCAGATGSGTACTTCCACAGAAAGATAGACTCTTTGAAAGAAAGATGGCACCAACT	1479
Db	1427	CCCTCGAATTCATATTTTGGCCCAAGAAAGATAGATTACTTTGAAGGAAACAAATCCAAACGTCCTT	1486
OY	1480	TTGTTTAGGAGAGAGAGACAAATGAGAGAGATGAGAGATTCAAGTGGAGATTCATATG	1539
Db	1487	TTTGTTAAAGACCGACGAGCTATGAAAGAGATATGAAAGCTTTAAAGTAGAGATTAATG	1546
OY	1540	CCTTAGTTGCCAAAGCCCAAGAAAGTTCCTGAAGAGATGACAAATCCAAATGAAACC	1559
Db	1547	CTCTTTGTTGGCCAAAGCACAGAAATATCCCTGAAGAAAGCTGGACAAATCAGATGCTACTC	1606
OY	1600	CCCTGGCCCTGGAACAAATGTTCTGTGATGATCCTTGAAATGATTCAGGTCTTCTTGCGCCAA	1659
Db	1607	CTGTGGCTGTGTAACAAACTAGAAATATCCTTGAAATGATACAGGTGTTCTTTCAGGCCATA	1666
OY	1660	GCGAGAGCCTTGACTGTGAGGGAAATGAACTGCCACAGATTGGTTTATGTTCTTAGAGAGA	1719
Db	1667	GTGGGGGCTGTGGAATACCGATGGAATATGAGCTGCCCTAATCTCATTTATGTTTCTGTGAAA	1726
OY	1720	AACGACCAAGGCTATTAACCATCATATAGAAAGCTGTGTGCTATGATGATGATTCAGTCCGAGTCT	1779
Db	1727	AGCGGCGCTGGATTTCAACACACAAAGAAAGGTGGAGCTATGAAATGACATTTGATTCGCTGTAT	1786
OY	1780	CTGCTGTACTTACCAATGCTTCATATTTGTTAACTTGATTTGATCATCATACATCAACA	1839
Db	1787	CTGCTGTCTTCTTACCAATGAGCAATATCTTTTGAACGTGGAATGATGATCATTTACTTTATATA	1846
OY	1840	ACACCAAGGCTATTAAGGAAGCAATGTGTTTATGAGAGCCCTTACTAGAGAAAGAAAG	1859
Db	1847	ACAGTAAAGGCTATTAAGAAAGCTATATGTGTTCAATGAATGAGACCCGGCTATTTGAAAGAAAGT	1906
OY	1900	TTTGTCTATGTATCAGATTCCTCTCAAGATTTTGAATGGATTTGATCCGCTATGACCGATATGCTA	1959
Db	1907	GTGCTATATGTCCAATTCCTCTCAAGATTTTGAACGCTATTTGATTTGACAGATGATGATGCCA	1966
OY	1960	ACCGGAATGTTGTCTTTTGTGATATCAACATGAAAGGTTTGGATGTATTCAGGGTCCAA	2019
Db	1967	ACAGGAATATATGTCCTTTTTCGATTTATTAACATGAAGGGGTGGATGATATCCAGGGTCCAG	2026
OY	2020	TTTATATGTTGGATCAGATGATGCTATTTTGAAGGACGCAATATATGTTATGATATGCCCA	2079
Db	2027	TATATATGAGGTACTGAGTTGTGTTTAAATAGGCAAGGCTCATATATGAGGTATGATCTGTGTT	2086
OY	2080	AAACAAAGAAAGCCACATCAAGACTTGCACAACCTGAGCCCAAGTGGTGTCTTGTGCTGT	2139
Db	2087	TGAAGGAAGAAAGATTTTGAACCAAAATATTA-----TTGTCAAGA	2125
OY	2140	GCTGCTTTGGCAATATGAGAACCAAAAGAAAGACTTACCAAAACCCAAAACAGAGAAAGAAAGT	2199
Db	2126	GCTGTGGGGGTCAAGAGAAAGAAAGTAAAGTAGCAAA--GAAGTATATATACGAAAAGA	2182
OY	2200	TATTTATTTTTCAGAAAGAAAGAAACCAATTCCTCTGCATATGCTCTTGGTGAATTTGACG	2259
Db	2183	GGAGAGGCATCTCAACAGATGACTCCATATGCTTCCATTTTCAATATGAGAGACATTCATATG	2242
OY	2260	AAGCTGTCCAGAGAGCTGAGATATGAAAGAGCCGCTATTTGTAAATTCAAACAAAAATTAGAAA	2319
Db	2243	AGGGTTTTTGAAGGTTATGATGATGATGAGAGCTTATTTCTAATGTCCACAGAGAGTGTAGAGA	2302
OY	2320	AGAAATTTGGCCCAATCTTCTGTGTTTGTATCATTCACACACTTCTCGAAGATGTTGAAACCT	2379
Db	2303	AGCGTTTTTGGTCAAGTCCCGCTATTTATTTGGGCAACCTTCATATGAAACAAAGCGGCAATTC	2362
OY	2380	TGAAGAGTGCAGATCTCTGCTCTCTTTTGAAGAAAGCTATATCATGTCAATTAATGTTGTGTT	2439
Db	2363	CACCAACAAACCAATCCCGCTACTCTTCTGAAAGAGGCTATTCATGTTATTAAGCTGTGTT	2422





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 17:49:46 ; Search time 7630 Seconds

(without alignments)  
12537.669 Million cell updates/sec

Title: US-09-720-383C-9

Perfect score: 3336

Sequence: 1 cttctccctgcgtgcgcg.....ctgagtcgtcttggaata 3936

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_eetba:\*  
2: em\_eetbm:\*  
3: em\_eetbn:\*  
4: em\_eetbv:\*  
5: em\_eetbv:\*  
6: em\_eetbv:\*  
7: em\_eetbv:\*  
8: em\_eetbv:\*  
9: gb\_eet1:\*  
10: gb\_eet2:\*  
11: gb\_eet3:\*  
12: gb\_eet4:\*  
13: gb\_eet5:\*  
14: gb\_eet5:\*  
15: em\_eetfun:\*  
16: em\_eetfun:\*  
17: em\_gse\_inv:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pln:\*  
20: em\_gse\_vtc:\*  
21: em\_gse\_vtc:\*  
22: em\_gse\_vtc:\*  
23: em\_gse\_vtc:\*  
24: em\_gse\_vtc:\*  
25: em\_gse\_vtc:\*  
26: em\_gse\_vtc:\*  
27: em\_gse\_vtc:\*  
28: gb\_gse1:\*  
29: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869.8	98.3	3897	11 AY103655	AY103655 Zea mays
2	2591	65.8	3783	11 AY104730	AY104730 Zea mays
3	2225.2	56.5	3788	11 AY103701	AY103701 Zea mays
4	1342.8	34.1	3898	11 AY110415	AY110415 Zea mays

5	1289.4	32.8	3763	11 AY108113	AY108113 Zea mays
6	1212.4	30.8	3696	11 AY110079	AY110079 Zea mays
7	1209.6	30.7	3728	11 AY112236	AY112236 Zea mays
8	1161.6	29.5	2872	11 AY104236	AY104236 Zea mays
9	811	20.6	877	29 CC379931	CC379931 PHO50TB
10	731.4	18.6	758	14 CD441181	CD441181 EV01N053
11	704.6	17.9	856	14 CB633016	CB633016 OS1IB117
12	686.2	17.4	859	14 CB645001	CB645001 OSJNB06H
13	683.2	17.4	864	14 CB672056	CB672056 OS1IB01L
14	672.4	17.1	862	14 CD437450	CD437450 BU01N0501
15	670.6	17.0	900	29 CC365622	CC365622 PUHRS2TD
16	652	16.6	792	14 CB671068	CB671068 OSJNB04H
17	641.2	16.3	783	14 CB671503	CB671503 OSJNB05B
18	626.6	15.9	1022	29 BZ827769	BZ827769 PUH017TB
19	621.8	15.8	795	14 CB677472	CB677472 OSJNB14G
20	612.8	15.6	806	14 CB629332	CB629332 OS1IB05H
21	596.4	15.2	758	14 CB672249	CB672249 OSJNB06D
22	594.6	15.1	714	14 CB629928	CB629928 OS1IB061
23	580.6	14.8	657	9 AM037064	AM037064 614025D05
24	578.6	14.7	765	14 CB676377	CB676377 OSJNB12J
25	570.6	14.5	670	29 CC160666	CC160666 1910C05.9
26	566.6	14.4	827	14 CB677498	CB677498 OSJNB141
27	549	13.9	701	13 BQ803259	BQ803259 WHE2835_E
28	546.8	13.9	770	14 CB681901	CB681901 OSJNB07C
29	539.6	13.7	789	14 CB672865	CB672865 OSJNB07C
30	539	13.7	591	14 CD425359	CD425359 SA1.11.F0
31	537.8	13.7	930	29 CC354726	CC354726 PUH207B
32	537.6	13.7	635	9 AM146609	AM146609 614074B01
33	533.4	13.6	642	9 AM053172	AM053172 614072C08
34	532	13.5	660	9 A1621823	A1621823 486027A04
35	525.4	13.3	539	13 BU197567	BU197567 946162C06
36	524.8	13.3	745	14 CD444028	CD444028 BU01N0434
37	524.4	13.3	695	13 BU009466	BU009466 WHE3359_G
38	522.6	13.3	755	13 BQ484013	BQ484013 WHE3515_D
39	517.2	13.1	760	10 BF628750	BF628750 HVSMB000
40	517	13.1	693	10 AL503528	AL503528 AL503528
41	515	13.1	700	9 A126453	A126453 614072C08
42	510	13.0	515	9 A1920493	A1920493 614025D05
43	505.4	12.8	731	13 BQ294855	BQ294855 WHE2855_B
44	502.8	12.8	731	13 CA019440	CA019440 HV11P23T
45	501.8	12.7	619	14	

## ALIGNMENTS

RESULT 1  
AY103655  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AY103655 3897 bp mRNA linear HTC 16-OCT-2002  
Zea mays PC0096398 mRNA sequence.  
AY103655  
AY103655.1 GI:21206733  
HTC.  
Zea mays  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 3897)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitcaltt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M., and Tinsley,S.V.  
Maize Mapping Project/Duront Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 3897)  
Coe,R.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSU, maize-map.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the



Db 1813 CATATTTTAACTTGATGTGATCACTACATCAACAACAGAGGCTATAAGAGAG 1872  
Qy 1861 CAATGTGTTTATGATGAGACCTTACTAGGAAAAGAGTTTGCTATGATGATCCCTC 1920  
Db 1873 CAATGTGTTTATGATGAGACCTTACTAGGAAAAGAGTTTGCTATGATGATCCCTC 1932  
Qy 1921 AAAATTTGATGGAGTTGATGCGCATGACCGATATGCTAACCGAATGTTGCTTTTGG 1980  
Db 1933 AAAATTTGATGGAGTTGATGCGCATGACCGATATGCTAACCGAATGTTGCTTTTGG 1992  
Qy 1981 ATATCAATGAAAAGTTGATGATGATGAGGCTGCAATTTATGTTGCTGATGATG 2040  
Db 1993 ATATCAATGAAAAGTTGATGATGATGAGGCTGCAATTTATGTTGCTGATGATG 2052  
Qy 2041 TATTAGAAAGGAGGCAATATATGTTATGATGCCCCCAAAAACAAGAGCACAATAA 2100  
Db 2053 TATTAGAAAGGAGGCAATATATGTTATGATGCCCCCAAAAACAAGAGCACAATAA 2112  
Qy 2101 GGACTTGCACTGCTGCGCCAAAGTGTGCTTTGCTGCTGCTTTGGCAATAGAGAC 2160  
Db 2113 GGACTTGCACTGCTGCGCCAAAGTGTGCTTTGCTGCTTTGGCAATAGAGAC 2172  
Qy 2161 AAAAGAGACTACCAAAACCCAAACAGAGAAAAGTTATTTTTCAGAAAAGAG 2220  
Db 2173 AAAAGAGACTACCAAAACCCAAACAGAGAAAAGTTATTTTTCAGAAAAGAG 2232  
Qy 2221 AGAACCAATCCCTGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
Db 2233 AGAACCAATCCCTGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2292  
Qy 2281 ATGAAAAGGCGGATATGTAATCAACAAAATTAGAAAAGAAATTTGGCCAACTTCTG 2340  
Db 2293 ATGAAAAGGCGGATATGTAATCAACAAAATTAGAAAAGAAATTTGGCCAACTTCTG 2352  
Qy 2341 TTTTGTGTTACCTCACTTCTGAGAAATGTTGAAACCTTGAAGATGCAAGTCTGCTT 2400  
Db 2353 TTTTGTGTTACCTCACTTCTGAGAAATGTTGAAACCTTGAAGATGCAAGTCTGCTT 2412  
Qy 2401 CTCTTTTGAAGAAGCTATATAGTCACTTATGTTGCTATGAAAGCAAGCAACATGCGG 2460  
Db 2413 CTCTTTTGAAGAAGCTATATAGTCACTTATGTTGCTATGAAAGCAAGCAACATGCGG 2472  
Qy 2461 GAAAAGAGATTGGCTGATCTATGATGATGATGATGATGATGATGATGATGATGATG 2520  
Db 2473 GAAAAGAGATTGGCTGATCTATGATGATGATGATGATGATGATGATGATGATGATG 2532  
Qy 2521 TGCAATGTCATGTTGGCGGATCAATTTACTGCAATACCTTAAACGGATGCTTCAAGGTT 2580  
Db 2533 TGCAATGTCATGTTGGCGGATCAATTTACTGCAATACCTTAAACGGATGCTTCAAGGTT 2592  
Qy 2581 CTGACCTCTGCAATCTTTCAATGATGCTTTCACAGGCTGCTTGGTGGCTTCTTGGGCTTA 2640  
Db 2593 CTGACCTCTGCAATCTTTCAATGATGCTTTCACAGGCTGCTTGGTGGCTTCTTGGGCTTA 2652  
Qy 2641 TTGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2700  
Db 2653 TTGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2712  
Qy 2701 TTTTGAAGAAGTTTCTCAATCAATCTTCTTGGTATGAGGATGATGATGATGATGATGATG 2760  
Db 2713 TTTTGAAGAAGTTTCTCAATCAATCTTCTTGGTATGAGGATGATGATGATGATGATGATG 2772  
Qy 2761 TGGCTTACTGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
Db 2773 TGGCTTACTGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2832  
Qy 2821 TGAATATATGTTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880  
Db 2833 TGAATATATGTTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2892  
Qy 2881 TCCTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Db 2893 TCCTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2952

Qy 2941 GGGTCATTGAGAGGTGTGCTCAACACCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000  
Db 2953 GGGTCATTGAGAGGTGTGCTCAACACCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3012  
Qy 3001 TAGCTGTGTGATGATACAGCTTCAACCGTGAATCAATCAAAAGGCTGAGATGATGAGAGTTCT 3060  
Db 3013 TAGCTGTGTGATGATACAGCTTCAACCGTGAATCAATCAAAAGGCTGAGATGATGAGAGTTCT 3072  
Qy 3061 CAGAGCTATATACATTTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
Db 3073 CAGAGCTATATACATTTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 3132  
Qy 3121 TGAATCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
Db 3133 TGAATCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3192  
Qy 3181 GGGGCCCCCTCTTGGAGAGCTATTTCTTGCATTTTGGGATGATGATGATGATGATGATGATG 3240  
Db 3193 GGGGCCCCCTCTTGGAGAGCTATTTCTTGCATTTTGGGATGATGATGATGATGATGATGATG 3252  
Qy 3241 TTCTCAAAAGGTTGTTGGAAGGCAAAACAGACACCAACGATGATGATGATGATGATGATG 3300  
Db 3253 TTCTCAAAAGGTTGTTGGAAGGCAAAACAGACACCAACGATGATGATGATGATGATGATG 3312  
Qy 3301 TTCTGCTGCTGCTCAATCTTCTGCTGCTTGGGTTGGATTTGATCTTCTTCTTGGAGAG 3360  
Db 3313 TTCTGCTGCTGCTCAATCTTCTGCTGCTTGGGTTGGATTTGATCTTCTTCTTGGAGAG 3372  
Qy 3361 ATGATGCTGCTGCTTCTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 3420  
Db 3373 ATGATGCTGCTGCTTCTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 3432  
Qy 3421 CTCGCCCAATCTGCAATGCTTGAAGATATTTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCT 3480  
Db 3433 CTCGCCCAATCTGCAATGCTTGAAGATATTTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCT 3492  
Qy 3481 CTGTAAGTAAAGACATGAAATGTCCCAAGTTCTTGTATCCATGATGATGATGATGATGATG 3540  
Db 3493 CTGTAAGTAAAGACATGAAATGTCCCAAGTTCTTGTATCCATGATGATGATGATGATGATG 3552  
Qy 3541 TATCTGAGAGATTAATCTGCGGAGAAATGAGAGGCTGCGGCAATCTTGTGATGATGATGATG 3600  
Db 3553 TATCTGAGAGATTAATCTGCGGAGAAATGAGAGGCTGCGGCAATCTTGTGATGATGATGATG 3612  
Qy 3601 TGAATACACCATATGCAATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 3660  
Db 3613 TGAATACACCATATGCAATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 3672  
Qy 3661 TAGATGGGCTGAGCCGAAACGCAAGTATTTTGAATCTGCACTGCTCCGCTGTAACAACT 3720  
Db 3673 TAGATGGGCTGAGCCGAAACGCAAGTATTTTGAATCTGCACTGCTCCGCTGTAACAACT 3732  
Qy 3721 TGGTCTCAATTAAGGAGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3780  
Db 3733 TGGTCTCAATTAAGGAGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3792  
Qy 3781 ATTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3840  
Db 3793 ATTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3852  
Qy 3841 AACAGAAATTAATGATTAATCTATGATTAATTAAGTATGATTAAT 3885  
Db 3853 AACAGAAATTAATGATTAATCTATGATTAATTAAGTATGATTAAT 3897

RESULT 2  
AY104730 3783 bp mRNA 11linear HTC 16-OCT-2002  
LOCUS Zea mays PC0100501 mRNA sequence.  
DEFINITION AY104730  
ACCESSION AY104730  
VERSION AY104730.1 GI:21207808  
KEYWORDS HTC.

SOURCE	ORGANISM	TITLE	JOURNAL REFERENCE	AUTHORS
Zea mays	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, W., Morgante, M. and Tingey, S.V. Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes Unpublished (2002) 2 (bases 1 to 3783)	Coe, E.H. Direct Submission Submitted (25-Apr-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schmale, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.	
FEATURES	LOCATION/Qualifiers	BASE COUNT	ORIGIN	Query Match
source	1..3783 /organism="Zea mays" /mol_type="rRNA" /db_xref="MaizEDB:635765" /db_xref="taxon:4577" /clone_lib="Maize Mapping Project/DuPont Cornsensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public configs to seed DuPont configs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"	933 a 908 c 1002 g 940 t	65.8%; Score 2591; DB 11; Length 3783; Best Local Similarity: 86.5%; Pred. No. 0; Matches 2884; Conservative 0; Mismatches 440; Indels 9; Gaps 2;	
QY	104 CTCTGCGCTTGTGGCGCGCGAGTAGTGAGGGAGGAACGATGAGGCGCGCGCT 163			
Db	146 CGTTGCCCGCCGGGATCTGGAGATCTGGTAGGCCAAGGGGATGAGGCGCAGCGCGGCT 205			
QY	164 GTGGCGCGGCTCCCAACCCGACAAGACTGCTCTCATCCGCGCGCACGGCGATCCGG 223			
Db	206 GGTCGCCCGGCTGCACAACCCGAAAGAGCTCCTCTCATCCGCGCGCGATGCGAGCCAG 265			
QY	224 GCCGAAGCGCGCGGGAGCAAGGGCAGGTGTGCCAGATTGGCGGCGAACGCTGG 283			
Db	266 GCCGAAGCCCATGACCAGCGGAACGGCCAGGTGTGCCAGATTGGCGGCGAGCGTGG 325			
QY	284 CTTTCCCCCGCGCGGACCCCTTGTGTGGCGTGCACAGATGCGCTTCCCCTCTGGCG 343			
Db	326 GCGCAACCCCGACGGGAGCGCTTGTGTGGCGTGCACAGATGCGCTTCCCCTCTGGCG 385			
QY	344 GGACTGCTACGAATAACGAGCGCGGAGAGGCGACGAAATGCGCCCGCAGTGCAGAATCG 403			
Db	386 GGACTGCTACGAGTAACGAGCGCGGAGAGGCGACGAAATGCGCCCGCAGTGCAGAATCG 445			
QY	404 ATACGAAGCGCGTCAAGGGCTGCGCAACGTGTGACCGGTGACGAGAGGAGAGACGGGTGGA 463			
Db	446 CTTCAAGGCGCTCAAGGGGTGCGCGCGGTGCGCGGAGCGAGAGAGAGAGCGGCTGGA 505			
QY	464 TGACCTGAGCAAACGAGTTCAACTGG---GACGCGCATGACTGCGAGTGTGTGGCCGAGTTC 520			
Db	506 CGACCTGGAGAAACGAGTTCAACTGGAGGAGCAAGACGACTCCAGTACTCTGCGGAGTTC 565			
QY	521 CATGCTTCAAGCCCAATGAGCTTACGCGCGGTGAGAGTGAACCTTAATGGCGGCCCAACAG 580			

Db	566	CATGCTCCAGCCCAACATGAGCTACGGCCGGGCGGCACTCGACGGCGTCCGCAAGC	625
Oy	581	TTTCAGCTCAACCCCAATGTTCCACTCTTCAACCAAGGGCAATGATGACATCCC	640
Db	626	ATTCCACCCCAATGTTCCCTCTCAACCAAGGACAGATGATGACATCCC	685
Oy	641	ACCGGACGACGACCGCGTGTGCTTCTTTCATGAGGTGTGTGGGGAAGAAGATCATCC	700
Db	686	GCCCGACACGACCCCTTGTGCTCTGTTCGTGTGGGTGGCGGGGGAAGAGATTCACCC	745
Oy	701	CCTTCCTTATTCGGATCCAGCTTACCTGTGCAACCAAGGCTATGAGCCATCCAAAGA	760
Db	746	TCTCCCGTAGCGGATCCCAACCTTCTGTGTGCAACCAAGGTCTATGAGCCCTTCMAAGA	805
Oy	761	TCTTGTGCATATGAGGTATGATGTGTGCTTGGAAAGAACGATGGAAATTGGAAGCA	820
Db	806	TCTCGCGCATATGAGCTACGGGAACGTATGAGAAAGAGATGAGAGCTGGAAGCA	865
Oy	821	GAGCAAGAGAGATGACACGACGGGAAATGATGTGTGTGTATGATGTGACATGTC	880
Db	866	GAAACAGAGAGGATGACACGAGAGAAACGATGGCGCGCATATATGTGTATGATGTC	925
Oy	881	TGATCTACCATATATGATATGAGGACAGCAAACTGTCCAGAAAATTCACATTCATC	940
Db	926	AGATCTACCATATATGATATGAGGCTTACACGCCATGTCCAGAAAATCCCGCTTCTTC	985
Oy	941	AAGCAGATTAATCCATATAGATGATGATTAATCATTTCCGCTTGTGTGTGGGATCTT	1000
Db	986	AAGCAATATCAACCCCTATAGAGATTAATTAATTCGCTATGTGTTTGTGTCTT	1045
Oy	1001	CTTCCACTACCGAGTATGATCCGGGTGAATGATGCAATTTGCTTGTGTGCTCATATCTGT	1060
Db	1046	CTTCCACTACCGAGTATGATCCGGGTGACGATGCAATTTGCTTATATGCTCATATCTGT	1105
Oy	1061	TATCGTGAATATCGTGTGCTGATGTCTGATCTTGAATTCATATCCCAAGTGTCCC	1120
Db	1106	GATCTGTGAATTTGGTTTGTGCATGTCTTGAATTTCTGACCAAGTTTCMAAGTGTCTCC	1165
Oy	1121	TATTGAGAGAGACTTACCTTGAACCGGCTGTCACTGAGGTTTCGACAAAGAAAGCCAGCC	1180
Db	1166	TATGAGAGAGAAACCTATCTTGAACCGGCTGATGATTAAGTTTGAACAAAGAAAGGCAATCC	1225
Oy	1181	ATCTCAATCTTGCTCCAAATGATTTCTTGTGTGATGACGGTGTATCCCTTAAGAAACCTCC	1240
Db	1226	TTCTCAATCCCGCTGTGATTTCTTGTGTGATGACGGTGTATCCCTTGAAGAAACCTCC	1285
Oy	1241	TTTGCTACCAAAATCTGTCTATCATCATCTTCCGTGTGATTAATCCTGTGTGAATAGCT	1300
Db	1286	ATTTGTCTCTTAATATCTTCTATCATCTTTCGTTGGTGAATTAATCAAGTTGATTAAGGT	1345
Oy	1301	TTCTTGTATTTCTGATGATGTGTGCTGATCTTAACTTTGAAGCATTAATCTGAAGC	1360
Db	1346	TTCAATGTCAAGTTTCTGATGATGTGTGCTGCAATCTGACATTTGAAGCATTTGTCTGAAGC	1405
Oy	1361	ATCTGAATTTGCAAAAGAAATGGGTTCTTTCTGCAACGATGACAAATATGAACCTCGCGC	1420
Db	1406	ATCTGAATTTGCAAAAGAAATGGGTTCTTTCTGCAAAAGATATAGCTTGAAGCTCGTGC	1465
Oy	1421	TCCAGAGTGTATTTCTCAACAGAAATGATTAATTTGAAGAAAGTGTGGAGCAAACTT	1480
Db	1466	TCCAGAGTGTATTTCTCAACAGAAATGATTAATTTGAAGAAAGTGTGGAGCAAACTT	1525
Oy	1481	TGTTAGGAGAGAGAGCAATGAAGAGAGATATGAGAAATTCAGAGTGAATATCAATGC	1540
Db	1526	TGTTAGGAGAGAGAGCAATGAAGAGAGATATGAGAAATTCAGAGTGAATATCAATGC	1585
Oy	1541	CTTAGTTCCAAAGCCCAAGAAATTTCTTGAAGAAAGATGACAAATGACAGATGAAACCC	1600
Db	1586	CTTAGTTCCTAAAGCCCAAGAAATTTCTTGAAGAAAGATGACAAATGACAGATGAAACCC	1645
Oy	1601	CTAGGCTTGAAGCAATGTTGTGTATCATCTTGAAGATGATTAAGATCTTCTTGTGC	1660

Db 1646 ATGGCCGGAATATATGTCCTGATCATCTGGAAATGATTCAGGTTTTCTTGTCMAAG 1705  
 Oy 1661 CGAGAGCCTTGACTGTCGGGGGAAATGAACTGCCAGATTGGTTTATGTTCTAGAGAA 1720  
 Db 1706 TGGTGGCCATATGTCGAAAGAAATGAGCTGCTGATTTGGTTTATGTTCTAGAGAAA 1765  
 Oy 1721 ACGACAGGCTATTAACATCATATAGAAAAGTGGTCTATGATGATGATTTGGTCCAGTCTC 1780  
 Db 1766 ACGGCAAGGCTATTAACATCATAGAAAGGCTGGTCTATGATGATGATTTGGTCCAGTCTC 1825  
 Oy 1781 TGGTGTACTTAACAAATGCTCATATTTGTTAACTTGGATTTGATGATCATCATCAACA 1840  
 Db 1826 TGGTGTACTTAACAAATGCTCATATTTGCTGAACTTGGATTTGATGATCATCATCAACA 1885  
 Oy 1841 CAGCAAGGCTATTAAGAAAGCAAGTGGTTTATGATGAGCCCTTACTAGAAAGAAAGGT 1900  
 Db 1886 TAGTAAGGCTATTAAGAAAGCAAGTGGTTTATGATGATGATTTGCTTGGAAATTAAGT 1945  
 Oy 1901 TTGCTATGATAGTTCCCTCAAGATTTGATGGAGTTGATGCGCATGACCGATATGCTTA 1960  
 Db 1946 TTGCTATGATGAGTTTCTCTCAAGATTTGATGGATTTGATGCGCATGATGATGCTTA 2005  
 Oy 1961 CCGGAATGTTGCTTTTGTATATCAATGAAAGTTGATGGATTTCAAGGTCAT 2020  
 Db 2006 CAGAAATGTTGCTTTTGTATATCAATGAAAGTTGATGGATTTCAAGGTCAT 2065  
 Oy 2021 TTATGTTGTTACTGATGATGATTTAGAAAGGAGGCACTTATATGTTTATGATGATGATG 2080  
 Db 2066 TTATGTTGTTACTGATGATGATTTAGAAAGGAGGCACTTATATGTTTATGATGATGATG 2125  
 Oy 2081 AACAAAGAGCCACATCAAGAGACTTGCATGCTGCGCCAAAGTGGTCTTTGCTGTG 2140  
 Db 2126 AACAAAGAGCCACATCAAGAGACTTGCATGCTGCGCCAAAGTGGTCTTTGCTGTG 2185  
 Oy 2141 CTGCTTTGCAATAGAAAGCAAGAACTAACAAACCCAAACAGAGA-----AGAA 2194  
 Db 2186 CTGTTTGGTAAACAGAAAGCAAGAAAGCAAGAACTTAAACCTTAATTTGAGAAA 2245  
 Oy 2195 AAAGTATATTTTTCAGAAAGAAAGAAACCAATCCCTGATATGCTCTTGGTGAAT 2254  
 Db 2246 GATTAAGAACTTTTAAAGAAAGAAAGAAATCAACCCCTGATATGCTCTTGGTGAAT 2305  
 Oy 2255 TGACCAAGCTGCTCCAGAGAGCTGAGATGAAAGGCGGATTTGTAATCAACAATAAT 2314  
 Db 2306 TGAATAGCTGCTCCAGAGAGCTGAGAAATGAAAGGCTGATTTGTAATCAACAATAAT 2365  
 Oy 2315 AGAAAGAAATTTGCGCAATCTTCTGTTTGTATCATTCACACTTCTCGAGATGTGG 2374  
 Db 2366 GGAAGAAATTTGCGCAATCTTCTGTTTGTATCATTCACACTTCTCGAGATGTGG 2425  
 Oy 2375 AACCTTGAAGAGTCAGATCCCTGCTTCTTTTGAAGAAAGCTATATCATATGATG 2434  
 Db 2426 AACCTTGAAGAGTCAGATCCCTGCTTCTTTTGAAGAAAGCTATATCATATGATG 2485  
 Oy 2435 TGGTATGAAAGCAAGAGTCAGAGGAAAGAGATGGCTGATCTATGATGATGATG 2494  
 Db 2486 TGGATATGAAAGCAAGAGTCAGAGGAAAGAGATTTGGTTGATTTATGATGATGATG 2545  
 Oy 2495 AGAAGATATTTCTACTGTTTCAAGATGATGATGATGATGATGATGATGATGATG 2554  
 Db 2546 AGAAGATATTTCTACTGTTTCAAGATGATGATGATGATGATGATGATGATGATG 2605  
 Oy 2555 AACCTTGAAGAGTCAGATCCCTGCTTCTTTTGAAGAAAGCTATATCATATGATG 2614  
 Db 2606 AACCTTGAAGAGTCAGATCCCTGCTTCTTTTGAAGAAAGCTATATCATATGATG 2665  
 Oy 2615 GATGCTGAGTGGGCTTTGGGCTATTTGAGATCTTCTTCAAGAAATGATGATGATG 2674  
 Db 2666 GATGCTGAGTGGGCTTTGGGCTATTTGAGATCTTCTTCAAGAAATGATGATGATG 2725  
 Oy 2675 GATGCTGAGTGGGCTTTGGGCTATTTGAGATCTTCTTCAAGAAATGATGATGATG 2734  
 Db 2726 GATGCTGAGTGGGCTTTGGGCTATTTGAGATCTTCTTCAAGAAATGATGATGATG 2785

Oy 2735 GATGCTGAGTGGGCTTTGGGCTATTTGAGATCTTCTTCAAGAAATGATGATGATG 2794  
 Db 2786 ATACCTTTGAGATCTATCCCGCTCTTGAGCTATTTGACATTTGCTGCTGCTGCT 2845  
 Oy 2795 GACAGGAAATTTATCACTCCAGAGCTGAATATATGTTTGGCAGCTGTTGATGATG 2854  
 Db 2846 GACAGGAAATTTATCACTCCAGAGCTGAATATATGTTTGGCAGCTGTTGATGATG 2905  
 Oy 2855 TTTTATCTGATTTTCTGATGAGATCTTGAATATGATGATGATGATGATGATGATG 2914  
 Db 2906 TTTTATCTGATTTTCTGATGAGATCTTGAATATGATGATGATGATGATGATGATG 2965  
 Oy 2915 TGACTGAGAGATGAGAGCTTCTGAGGCTATTTGAGATGATGATGATGATGATGATG 2974  
 Db 2966 TGACTGAGAGATGAGAGCTTCTGAGGCTATTTGAGATGATGATGATGATGATGATG 3025  
 Oy 2975 TGTGTTCCAGAGACTTCTCAAGGCTATGATGATGATGATGATGATGATGATGATG 3034  
 Db 3026 TGTGTTCCAGAGACTTCTCAAGGCTATGATGATGATGATGATGATGATGATGATG 3085  
 Oy 3035 AAAGGAGAGATGATGAGAGCTTCAAGGCTATGATGATGATGATGATGATGATG 3094  
 Db 3086 CAAGGAGAGATGATGAGAGCTTCAAGGCTATGATGATGATGATGATGATGATGATG 3145  
 Oy 3095 GATACCTCTTCACTTCTGCTTCTTATGATGATGATGATGATGATGATGATGATG 3154  
 Db 3146 GATACCTCTTCACTTCTGCTTCTTATGATGATGATGATGATGATGATGATGATG 3205  
 Oy 3155 TGGATCAATTAAGGATATGATGATGATGATGATGATGATGATGATGATGATG 3214  
 Db 3206 TGGATCAATTAAGGATATGATGATGATGATGATGATGATGATGATGATGATG 3265  
 Oy 3215 TTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3274  
 Db 3266 TTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3325  
 Oy 3275 ACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3334  
 Db 3326 GCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3385  
 Oy 3335 TCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3394  
 Db 3386 CCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3445  
 Oy 3395 TTGCACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3427  
 Db 3446 TTGCACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3478

RESULT 3  
 AT103701  
 LOCUS 3788 bp mRNA linear HTC 16-OCT-2002  
 DEFINITION Zea mays P00120363 mRNA sequence.  
 ACCESSION AY103701  
 VERSION AY103701.1 GI:21206779  
 KEYWORDS HTC.  
 SOURCES Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 3788) Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., White, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.  
 Maize Mapping Project/Dupont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 3788)  
 REFERENCES  
 JOURNAL  
 TITLE Direct Submission  
 AUTHORS Coe, E.H.  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA





Db 1911 ATCAACAAATAGCAAGCCATTAAGAGGCTATGTGTTTCATGATGATCCTTTGGTGGGG 1970  
 Qy 1893 AAGAAAGTTTGTATGTATGACAGTTCCCTCAAAAGATTGATGAGATGACCAATACGA 1952  
 Db 1971 AAGAAAGTGTGTATGTATGACAGTTCCCTCAAAAGATTGATGAGATGACCAAAATATGCA 2030  
 Qy 1953 TATGCTAACCCGAGATGTTGTTTGTATGATCAATGAAAGTTTGGATGGTATTCAG 2012  
 Db 2031 TAGGCTAACAGGAGATGTTGTTTGTATGATCAATGAAAGTTTGGATGGTATTCAG 2090  
 Qy 2013 GGTCAATTTATGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2072  
 Db 2091 GACACCATTTATGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2150  
 Qy 2073 GCGCCCAAAACAAAGAGCCACCATCAAGAGCTTGCACTGCTGGCCCAAGTGTCTT 2132  
 Db 2151 GCTCTTAAACCAAGAGCCACCATCAAGAGCTTGCACTGCTGGCCCAAGTGTCTT 2210  
 Qy 2133 TGCTGTGTGTGTTTGGCAATGGAAGCAAAAGAGCTTACCAACCCCAAGAGAG 2192  
 Db 2211 TCTTCTGTGTGCAAGAGCAAGATTAAGAGAGCTTACCAACCCCAAGAGAG 2270  
 Qy 2193 AAAAGTTATTTTCAAGAAAGAGAGCAATCCCTGCAATATGCTTGGTGA 2252  
 Db 2271 AAGAAAGTTATTTTCAAGAAAGAGAGCAATCCCTGCAATATGCTTGGTGA 2330  
 Qy 2253 ATTGAAGAGCTGCTCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312  
 Db 2331 ATTGAAGAGCTGCTCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2390  
 Qy 2313 TTAGAAAGAAATTTGGCAATCTTCTGTTTGTATGATGATGATGATGATGATGATGAT 2372  
 Db 2391 CTAGAGAGAAATTTGGCAATCTTCTGTTTGTATGATGATGATGATGATGATGATGAT 2450  
 Qy 2373 GGAACCTTGAAG 2432  
 Db 2451 GGAACCTTGAAG 2510  
 Qy 2433 TGTGTTATGAAG 2492  
 Db 2511 TGTGTTATGAAG 2570  
 Qy 2493 AAGAAAGATTTCTAAG 2552  
 Db 2571 AAGAAAGATTTCTAAG 2630  
 Qy 2553 ATACCTAAAG 2612  
 Db 2631 ATCCGAG 2690  
 Qy 2613 CAGGTGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGG 2671  
 Db 2691 CAGGTGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTTGGTGG 2750  
 Qy 2671 TTGTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2731  
 Db 2751 TTGTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2810  
 Qy 2731 CGTGTATCTTGAAG 2791  
 Db 2811 CGTGTATCTTGAAG 2870  
 Qy 2791 ATTGAAG 2851  
 Db 2871 GCTACAGGGAG 2930  
 Qy 2852 ACTTTTATCTGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2911  
 Db 2931 ACTTTTATCTGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2990  
 Qy 2912 TGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2971  
 Db 2991 CGACGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3050

Qy 2972 TGCTGTGTGTGAG 3031  
 Db 3051 CGCGGTGTGTGAG 3110  
 Qy 3032 ATCAAG 3091  
 Db 3111 GTGAG 3170  
 Qy 3092 ATTGATATCTTCAACAG 3151  
 Db 3171 GCTGATATCTTCAACAG 3230  
 Qy 3152 AATGAG 3211  
 Db 3231 GAG 3290  
 Qy 3212 ATTTGGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3271  
 Db 3291 GTTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3350  
 Qy 3272 GACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3331  
 Db 3351 GAGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3410  
 Qy 3332 GGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3391  
 Db 3411 GGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3470  
 Qy 3392 GGAATGCAACT 3402  
 Db 3471 GGAATGCAACT 3481

**RESULT 4**  
**AY110415**  
**LOCUS** AY110415 3898 bp mRNA linear HTC 17-OCT-2002  
**DEFINITION** Zea mays CL1166\_1 mRNA sequence.  
**ACCESSION** AY110415  
**VERSION** AY110415.1 GI:21214824  
**KEYWORDS** HTC.  
**SOURCE** Zea mays  
**ORGANISM** Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
**REFERENCE** Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgente, M. and Tingey, S.V.  
**AUTHORS** Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
**TITLE** Unpublished (2002)  
**JOURNAL** 2 (bases 1 to 3898)  
**REFERENCE** Coe, B. H.  
**AUTHORS** Direct Submission  
**TITLE** Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
**COMMENT** If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, maize-map.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.  
**FEATURES** Location/Qualifiers  
**source** 1..3898  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="MaizeDB:630057"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST



[illegible]

QY	3208	TTGATTTGGGTGATTTGTGCATCTTTATCCCTTTCTCAAGGTTGGTGGAAAGGCMAA	3267
Db	3220	TTTGATCTGGGTGATTTCTCATCTTGACAGTTTCTCAAGGGCTCATGGGAAGGACGA	3279
QY	3268	ACAGGACACCAACATGTGATCGTGTGGTGCATTTGATCGAGCTTCATCTTCGCTCC	3327
Db	3280	ACCGCACACCAACATGTCATGTCTGTGGTGCATCTTTGACATTACTTCTCTTGC	3339
QY	3328	TTTGGGTTGGATTCATCTTTCTCT	3352
Db	3340	TGTGGGTGAAGATCGATCTTTTCA	3364

RESULT 5					
AY108113	AY108113	3763 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	AY108113				
DEFINITION	Zea mays	PC0126465	mRNA	sequence.	
ACCESSION	AY108113				
VERSION	AY108113.1	GI:21211191			
KEYWORDS	HTC				
SOURCE	Zea mays				

REFERENCE AUTHORS	TITLE	JOURNAL REFERENCES
1 (bases 1 to 3763) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitelc,t,M.S., Arthur,L.W., Hanafey,M., Morgane,M. and Tingey,S.V.	Maize Mapping Project/DuPont Consensus Sequences for Design of Ove,rge Probes	Unpublished (2002)
2 (bases 1 to 3763)		

**JOURNAL COMMENT**

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, [maizegap.org](http://maizegap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR, [www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES	SOURCE
location/Qualifiers	
1. .3763	
/organism="Zea mays"	
/mol_type="mRNA"	
/db_xref="MaizEDB:637367"	
/db_xref="taxon:4577"	
/clone_id="Maize Mapping Project/DuPont Cornensus Library"	
/note="this sequence is part of a project of BSR assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"	
863 a 984 c 1011 g 904 t 1 others	

Qy	210	GAGCGGCATCCCGGCGCGAAGCCCGCGGAGAGCAGAA	CGGCGAGCGTGTCCAGATTTC	269
Dy	344	GGCACC CGGACCGCGCTGAACTCGGAGAGCGCGGTGTG	CGGACAGAGTGTCCAGATTTC	403
Qy	270	GGCAGCAGCGTGGCGCTTGGCCCCCGCGGAGAACCCCTT	CGTGGCGTCAGACGAGTGCCTC	329
Dy	404	GGCAGCAGCGTGGCGCACACGCGGAGAGGGAGACGTTCTT	CGCGCTTCGACAGTCTGGGG	463

Query Match 32.8%; Score 1289.4; DB 11; Length 3763;  
 Best Local Similarity 64.7%; Pred. No. 1.2e-277;  
 Matches 2112; Conservative 1; Mismatches 1042; Indels 108; Gaps 9;







Db 2801 ATCAACACAAATTATTCACATCATATCCGCTCTCTGACTGATATTCACA 2860  
 Qy 2781 GCCATCTGTTATTTAGCAGGAAATTTATCATCCAGAGCTGAATATTCACCCG 2840  
 Db 2861 GGAATTTGCTCTCTCACTGGAAATTCATATCCCAAAATAGTAACTTAAGAGTGT 2920  
 Qy 2841 TGGTTCATGTCATTTTATTCATGATTTTGTGACAGCATCTTGAATATGATGAGT 2900  
 Db 2921 TGGTTCATGTCATTTTATTCATGATTTTGTGACAGCATCTTGAATATGAGT 2980  
 Qy 2901 GGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2960  
 Db 2981 NNN 3040  
 Qy 2961 TCACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3020  
 Db 3041 GCGCATTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3100  
 Qy 3021 TTCACCTGTCATCAAAAG--GTGAGATGATGAGAGTCTCAGAGCTATATATTC 3077  
 Db 3101 TTCACCTGTCATCAAAAG--GTGAGATGATGAGAGTCTCAGAGCTATATATTC 3160  
 Qy 3078 AAATGACTACTTATATGATATCTTACACCTGCTGCTGCTGCTGCTGCTGCTGCTG 3137  
 Db 3161 AAGTGAACACGCTCTGATCCCAACCACTATTTGATCATCACTGCTGCTGCTGCTG 3220  
 Qy 3138 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3197  
 Db 3221 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3280  
 Qy 3198 AAGTATTTCTTGTGATTTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3257  
 Db 3281 AAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3340  
 Qy 3258 GGAAGGCAAAACAGACACCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3317  
 Db 3341 GGAAGGCAAAACAGACACCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3400  
 Qy 3318 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3377  
 Db 3401 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3460  
 Qy 3378 GAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3411  
 Db 3461 GCGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3494  
 RESULT 7  
 LOCUS AY112236 3728 bp mRNA linear HTC 17-OCT-2002  
 DEFINITION Zea mays C11160\_1 mRNA sequence.  
 ACCESSION AY112236  
 VERSION AY112236.1 GI:21216826  
 KEYWORDS HTC.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 BUKRYOTCA, VIRIDIPANTAE, STREPTOPHYTA, EMBRYOPHYTA, TRACHEOPHYTA;  
 SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES; POACEAE; PACCAD  
 CLADE; PANICOLIDEAE; ANDROPOGONEAE; Zea.  
 1 (bases 1 to 3728)  
 HANEY, C.F., DOLAN, M., MIAO, G.H., VOGEL, J.M., WHITESIDE, M.S.,  
 ARTHUR, L.W., HANFELY, M., MORGANTE, M. and TINGEY, S.V.  
 Maize Mapping Project/Dupont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 3728)  
 Coe, E.H.  
 JOURNAL Direct Submission  
 REFERENCE Submitted (25-Apr-2002) Maize Mapping Project, University of  
 AUTHORS Missouri, Columbia, MO 65211, USA  
 TITLE If you are interested in getting corresponding physical clones,  
 JOURNAL these are publicly available from ZmDB and may be found by BLAST  
 COMMENT Searching at MSL, maize-map.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Maltot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.  
 Location/Qualifiers  
 FEATURES  
 Source  
 1. 3728  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="maizedb:630049"  
 /db\_xref="taxon:4577"  
 /clone\_id="Maize Mapping Project/Dupont Consensus  
 Library"  
 /note="this sequence is part of a project of BSR  
 assemblies resulting from the application of public  
 contigs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"  
 BASE COUNT 782 a 845 c 925 g 878 t 298 others  
 ORIGIN

Query Match 30.7%; Score 1209.6; DB 11; Length 3728;  
 Best Local Similarity 59.9%; Pred. No. 8.7e-260;  
 Matches 1963; Conservative 0; Mismatches 1210; Indels 102; Gaps 7;

Qy 210 GACGGCATCCCGGCGAAGCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269  
 Db 244 GAGAGCGGAG 303  
 Qy 270 GCGGAG 329  
 Db 304 GCGGAG 363  
 Qy 330 TTCCCGCTGCGGAG 389  
 Db 364 TTCCCGCTGCGGAG 423  
 Qy 390 CAGTGAAG 449  
 Db 424 NAGTGAAG 483  
 Qy 450 GAG 497  
 Db 484 GAG 543  
 Qy 498 GACTGAG 557  
 Db 544 CAG 603  
 Qy 558 GACCTATAG 617  
 Db 604 GATGTTGCGGCGGAG 663  
 Qy 618 GGGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677  
 Db 664 GAGATCCTCGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723  
 Qy 678 GGTGGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 734  
 Db 724 GGTGCTTCCCTGACATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 783  
 Qy 735 CCCAGTCTATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785  
 Db 784 TTTCCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843  
 Qy 786 GTTCTTGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 829  
 Db 844 GTTCTTGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903  
 Qy 830 -----GAGATGACACAG 873  
 Db 904 ATGACGATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 963





Db 3094 TTCACCGTCACCTCCAGGCGCTCGGATGAGAAGCGGCACTTCGCGAGCTGTACATGTTG 3153  
Oy 3078 AATGAGTACCTTATGATACCTCTCAACCTTGTCTTATGATCAATGATG 3137  
Db 3154 AAGTGAAGCACTTCTATCCGCCACACCACTCTGATCATTAACCTGGTCGGCTT 3213  
Oy 3138 GTGCTGGCGTTCAATGATGATCAATAACGATATGATGATGAGGCGCCCTCTTTGGG 3197  
Db 3214 GTTCCGCGCATCTCTACGCCATCAACAGCGGGTACGATCGTGGGGTCCGCTCTTTGGG 3273  
Oy 3198 AAGCTATCTTGTGATTTTGGGATTTGTCATCTTTATCCCTTTCTCAAAAGTTGGTT 3257  
Db 3274 AAGCTCTTCTGCGCTTCTGATGATCGTTACCTGTACCGCTTCTCAAGGGTCTCATG 3333  
Oy 3258 GAAAGGCAAAACAGAGACCAACGATGTCATCGTGTGTCATCTGCTGGCTCAATG 3317  
Db 3334 GGTGGCAAGAACCGACCCGACCATCGTGGTGTGCTGGGAGATCTCTGGCGTGCATG 3393  
Oy 3318 TTCTGCTCTTGTGGGTTGATGATGATCTTCTTCTTGGAAAGATGATGTCGGCTTCTT 3377  
Db 3394 TTCTGCTTGTGTGGGTTGATGATGATCGTTTCAACCAACCGGTGACTGGCCGGATCT 3453  
Oy 3378 GAGAGTGTGTTGGATGATGCACTAGATGTCAG 3412  
Db 3454 CGAAGTGTGGATCACTGCTAGGAGGTGAG 3488

RESULT 8  
AY104236  
LOCUS AY104236 2872 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PCO121439 mRNA sequence.  
ACCESSION AY104236  
VERSION AY104236.1 GI:21207314  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 2872)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 2872)  
Coe,E.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schubler, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.  
FEATURES  
Source  
1..2872  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizEDB:637064"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
configs to seed Dupont configs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

BASE COUNT 730 a 600 c 688 g 852 t 2 others  
ORIGIN

Query Match 29.5%; Score 1161.6; DB 11; Length 2872;  
Best Local Similarity 68.4%; Pred. No. 4,7e-249;  
Matches 1679; Conservative 1; Mismatches 740; Indels 34; Gaps 4;  
Oy 903 GCAGACAAACATGTCAGAGAAAATTCATCTTCATCAAGCCAGATTAATCATATAG 962  
Db 13 GCAGGCTACCTTAAGTCGATGTCATCCGATATCTCAAAAGAGCTTAACCTTATCG 72  
Oy 963 ATGATTAATCAATATTCGGCTTGTGGTGTGGGGTCTTCTCCACTACCGAGATGAT 1022  
Db 73 ATCGATATGTTCTCCGGCTTATCATCTATGTTCTTCTTCAATATCGATTAACAT 132  
Oy 1023 CCGGTGAATGATGATTTGCTTGTGGCTCATCTGTATCTGTATCTGTAAATCTGTTGCC 1082  
Db 133 CCAATGGAATGCTTATGATGTTGTGTGTGTGTATCTGTATTTGTGAAGTTGGTTGCC 192  
Oy 1083 ATGCTTGATTTCTTGATCAATTCCAAAGTGTTCCTTATGAGAGAGACTTACTTA 1142  
Db 193 TTGCTTGGCTTGAATCAAGTTCCAAAGTGTATCTTATCAACCGTGAACCTTACCTC 252  
Oy 1143 GACCGCTGTCACTGAGGTTGAGACAGAGAGCCAGCATCTCACT-TGCTCAATGA 1201  
Db 253 GATAGACTTGTGATGATATGATAGAGAGGAGGAGCATCCAGTTGGGCTCAATCGA 312  
Oy 1202 TTCTTTGTGAGTACGGTGTATCCCTTAAAGAACTCTTGTGTCAACAATATCTGT 1261  
Db 313 TGTCTTTGTATGATCAATGATGATCACTTAAGAACTCTCTTAATTAATGCGCAACTGT 372  
Oy 1262 TCTATCTATCTTTCGGTGTATATCTGTTGATAGGTTCTGTATGTTTGTGATGA 1321  
Db 373 CCTGTCATCTTCTGCTGTGATTAACCTGTTGAACAAAGTATCATGTTATGTTTGTGATGA 432  
Oy 1322 TGGTGTGCAATGCTTAACGTTTGAAGATTAATCTGAAACATCTGAATTTGCAAGAAATG 1381  
Db 433 CGGTTACGCTATGTTGACTTTTGAAGCGCTATCGAAACCGCAAGTTTGAAGAAATG 492  
Oy 1382 GGTTCCTTTGCGCAAGAGTACAAATATTTGAACCTCGGCTCGAGAGTATCTTCAACA 1441  
Db 493 GGTTCCTTTTGCAGAAACAAATATTTGAACCTGAGGCTCGAGAGTATCTTGTGCTCG 552  
Oy 1442 GAAGATGACTACTTGAAGAGCAAGTGGGACCAACTTTGTTAGGAGAGAGAGCAAT 1501  
Db 553 AAGATGATTAATCTTAAGAGCAAAATATCAACCTCTTTGTGAAGAGGCGGGCAT 612  
Oy 1502 GAAGAGAGATATAGAAATTTCAAGGTGAGATTAATGCTTACTTATGCCAAAGCCAGAA 1561  
Db 613 GAGAGGAGAGTGTGAAGATTCAAAGTACGATGATGATGCTTGTCCAAAGCGCAAAA 672  
Oy 1562 AGTTCCTGAAGAGATGAGCAATGCAAGATGGAACCCCTGGGCTGGAACCAATGTTCC 1621  
Db 673 AATACCTGAGAGGAGGCTGAGACATGCTGATGCACTCTTGGCTGGAAATTAACCTTAG 732  
Oy 1622 TGATCATCTGGAATGATGATGAGTCTTCTTGGCAAGCGAGGCTTGAAGGAGG 1681  
Db 733 AGATCATCTGAGATGATTCAGATATCTTGGGCAAGTGTGGGCTTGAACAGATGG 792  
Oy 1682 AATGAACTGCAAGATGTTGTTATGTTTCTAGAGAGAAAGACAGAGCTTAACATCA 1741  
Db 793 GAAAGATGTCAGAGGCTGTTGTTATGTTTCTCGGAGAAAGAGGCAAGGCTCCAGACCA 852  
Oy 1742 TAAAGAGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
Db 853 CAAGAGGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912  
Oy 1802 ATATTGTAAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1861  
Db 913 TTAATCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972  
Oy 1862 AATGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1921  
Db 973 TATGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1032



Db 763 GCATCCTAGAAATGAGAGAGAGTGTGTGGAATTGATGACTGTGGAGAAATGAGCACT 704

Qy 2938 TCTGGAGTCAATGGAGAGTGTGCTTCAACACCTCTTTGCTGTGTTCAGGAGACTTCTCAAG 2997

Db 703 TCTGGAGTCAATGGAGAGTGTGCTTCAACACCTCTTTGCTGTGTTCAGGAGACTTCTCAAG 644

Qy 2998 TCAATAGCTGTGTGATACAAAGCTTCAACCGTACATCAAAAGGTGAGATGATGAGAGT 3057

Db 643 TCATAGCTGTGTGTGATACAAAGCTTCAACCGTACATCAAAAGGTGAGATGATGAGAGT 584

Qy 3058 TCTCAGAGCTATATACATTCAAATGAGACTTATGATACCTCTTACACCTTGTCTTC 3117

Db 583 TCTCAGAGCTATATACATTCAAATGAGACTTATGATACCTCTTACACCTTGTCTTC 524

Qy 3118 TATTGAACCTTCAATGAGTGTGCTGCGGCTTCAAAATGATGATCAATTAACGATTAAGT 3177

Db 523 TATTGAACCTTCAATGAGTGTGCTGCGGCTTCAAAATGATGATCAATTAACGATTAAGT 464

Qy 3178 CATGGAGGCCCCCTCTTTGGGAAAGCTATCTTTGCAATTTGGGATGTGTCATCTTTATC 3237

Db 463 CATGGAGGCCCCCTCTTTGGGAAAGCTATCTTTGCAATTTGGGATGTGTCATCTTTATC 404

Qy 3238 CTTTCTCAAGGTTTGTGGAAGGCAAAACAGACCAACGATGTGTCATGCTGTGT 3297

Db 403 CTTTCTCAAGGTTTGTGGAAGGCAAAACAGACCAACGATGTGTCATGCTGTGT 344

Qy 3298 CCATCTGCTGAGCTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3357

Db 343 CCATCTGCTGAGCTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284

Qy 3358 AGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3417

Db 283 AGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224

Qy 3418 CAGCTCCCCCAATCTGCAATGCTTGAAGTATATTTCTGCTGCTGCTGCTGCTGCTGCT 3477

Db 223 CAGCTCCCCCAATCTGCAATGCTTGAAGTATATTTCTGCTGCTGCTGCTGCTGCTGCT 164

Qy 3478 TGTCTGTGATTAAGAGACATGAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3537

Db 163 TGTCTGTGATTAAGAGACATGAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104

Qy 3538 TAAATCTGAGAGATTAATCTGCGGAGAAATGAGAGCTGCGGCAATCTTGTGCAAGTGG 3597

Db 103 TAAATCTGAGAGATTAATCTGCGGAGAAATGAGAGCTGCGGCAATCTTGTGCAAGTGG 44

Qy 3598 CCGTGAATACAGCATATGCAAGTGTGATGTCGACGATTC 3640

Db 43 CCGTGAATACAGCATATGCAAGTGTGATGTCGACGATTC 1

RESULT 10

CD441181/c 758 bp mRNA linear EST 03-JUN-2003

LOCUS EL010553C04.5 Endospem\_5 Zea mays cDNA, mRNA sequence.

DEFINITION CD441181

ACCESSION CD441181

VERSION CD441181.1 GI:31356824

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 758)

La1,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and Messing,J.

Sequencing of the maize endospem ESTs

Unpublished

CONTACT: Lai, Jinseng

Dr. Joachim Messing's lab

Wakeman Institute, Rutgers University

190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801

FEATURES

source

location/Qualifiers

1..758

/organism="Zea mays"

/mol\_type="mRNA"

/culivar="W22"

/db\_xref="taxon:4577"

/issue\_type="Endospem of 7-21DAP"

/clone\_lib="Endospem\_5"

/note="Vector: pBluescript SK-, site\_1: EcoRI; site\_2: XhoI"

BASE COUNT 236 a 177 c 144 g 201 t

ORIGIN

Query Match 18.6%; Score 731.4; DB 14; Length 758;

Best Local Similarity 99.2%; Pred. No. 8.5e-153;

Matches 735; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3144 GGGGTTCAATGGATGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3203

Db 758 GGGGTTCAATGGATGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 699

Qy 3204 TTTTTCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3263

Db 698 TTTTTCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 639

Qy 3264 CAAAACAGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3323

Db 638 CAAAACAGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 579

Qy 3324 CTCTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3383

Db 578 CTCTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 519

Qy 3384 TGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3443

Db 518 TGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 459

Qy 3444 AAGTATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3503

Db 458 AAGTATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 399

Qy 3504 TCCCAAGTTTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3563

Db 398 TCCCAAGTTTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 339

Qy 3564 AAATGAGGCTGCGGCAATCTTGTGCAAGTGTGAGGCGGGAATCAGCATATGCAAGTGT 3623

Db 338 AAATGAGGCTGCGGCAATCTTGTGCAAGTGTGAGGCGGGAATCAGCATATGCAAGTGT 279

Qy 3624 TTGATGTCAGCATCTTTTATTAATCTGTCGATGATGATGATGATGATGATGATGATG 3683

Db 278 TTGATGTCAGCATCTTTTATTAATCTGTCGATGATGATGATGATGATGATGATGATG 219

Qy 3684 AGTATTTTGAATCTGCACTGCTCCCGTGAACAACTTGTCTCAATTAAGCAGGCGAG 3743

Db 218 AGTATTTTGAATCTGCACTGCTCCCGTGAACAACTTGTCTCAATTAAGCAGGCGAG 159

Qy 3744 AATGATCTGTCAGTGAAG 3803

Db 158 AATGATCTGTCAGTGAAG 99

Qy 3804 AGGCTGTGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3863

Db 98 AGGCTGTGTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 39

Qy 3864 TAGTTAATTAAGTATGATAA 3884

Db 38 TAGTTAATTAAGTATGATAA 18

RESULT 11  
 CB633016 856 bp mRNA linear EST 08-APR-2003  
 LOCUS OS11B1J121.f OS11B Oryza sativa (indica cultivar-group) cDNA  
 DEFINITION Oryza sativa (indica cultivar-group) cDNA  
 ACCESSION CB633016  
 VERSION CB633016.1 GI:29628005  
 KEYWORDS EST.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 856)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 JOURNAL Unpublished  
 COMMENT Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR Primers  
 FORWARD: gta aac cga cgg cca gtc  
 BACKWARD: gga aac agc tac gac cat g  
 Plate: 11 row: J column: 21  
 Seq primer: gta aac cga cgg cca gtc.  
 Location/Qualifiers  
 1..856  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="IR36"  
 /db\_xref="taxon:39946"  
 /clone="OS11B1J121"  
 /issue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH108"  
 /clone\_lib="OS11RB"  
 /note="Vector: pBluescript II KS +; Site 1: EcoRI, Site 2:  
 XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"  
 BASE COUNT 222 a 165 c 203 g 266 t  
 ORIGIN  
 Query Match 17.9%; Score 704.6; DB 14; Length 856;  
 Best Local Similarity 89.0%; Pred. No. 8.7e-147;  
 Matches 761; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
 OY 2219 AGAGAACCAATCCCTGATGCTCTTGGGAATTGACGAAGCTGCTCCAGAGCTGA 2278  
 DB 2 AGAAAAATCAATCCCTGATGCTCTTGGGAATTGACGAAGCTGCTCCAGAGCTGA 61  
 OY 2279 GAATGAAAGCCCGGATATGTAATCAACAAAATTAAGAAAGAAATTTGCCAATCTTC 2338  
 DB 62 AAATGAAAGGCTGTAATGTAATCAACGAAGCTGAAAAGAAATTTGCCAATCTTC 121  
 OY 2339 TGTCTTTGTATCACTCCACACTCTCCGAAATGATGGAACCTTGAAGAGTGCATCTGC 2398  
 DB 122 TGTCTTTGTGCAATCCCACTCTCTTGAAGATGGGAACCTTGAAGAGTGCATCTGC 181  
 OY 2399 TTTCTTTTGAAGAGCTATACATGATGATGATGATGATGATGATGATGATGATGATG 2458  
 DB 182 ATCTCTTAAAGAGAGATATTCATGCTATGCTATGCTATGCTATGCTATGCTATGCT 241  
 OY 2459 GGGAAAAGAGTGGCTGATCTATGATGATGATGATGATGATGATGATGATGATGATG 2518  
 DB 242 GGGAAAAGAGTGGCTGATCTATGATGATGATGATGATGATGATGATGATGATGATG 301  
 OY 2519 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578

DB 302 GATGACCTCCATGTTGGCGGTCAATTTACTGCACTCAATCACTGCACTTCAAG 361  
 OY 2579 TTCTGACCTCTGAATCTTTTCAAGATGCTTTTCAAGAGTGTCTTGGGCTCTTGGGCTC 2638  
 DB 362 TTCTGACCACTCAATCTTTTCAAGATGCTTTTCAAGAGTGTCTTGGGCTCTTGGGCTC 421  
 OY 2639 TATGAGATCTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2698  
 DB 422 TGTGAAATATCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 481  
 OY 2699 ATTTTGAAGATTTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 2758  
 DB 482 ATGTTGAAAGATTTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 541  
 OY 2759 CTGGGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2818  
 DB 542 CTGGGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601  
 OY 2819 GCTGAATATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2878  
 DB 602 GCTTGAATATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 661  
 OY 2879 CATCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2938  
 DB 662 CATCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721  
 OY 2939 CTGGGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2998  
 DB 722 TTGGGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781  
 OY 2999 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3058  
 DB 782 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841  
 OY 3059 CTCAGAGCTATATAC 3073  
 DB 842 CTCAGAGCTATATAC 856  
 RESULT 12  
 CB645001 859 bp mRNA linear EST 08-APR-2003  
 LOCUS OSJNB06H07.f OSJNB Oryza sativa (japonica cultivar-group) cDNA  
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA  
 ACCESSION CB645001  
 VERSION CB645001.1 GI:29639992  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 859)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 JOURNAL Unpublished  
 COMMENT Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR Primers  
 FORWARD: gta aac cga cgg cca gtc  
 BACKWARD: gga aac agc tac gac cat g  
 Plate: 06 row: H column: 07  
 Seq primer: gta aac cga cgg cca gtc.  
 Location/Qualifiers



[illegible]

FEATURES		Location/Qualifiers
SOURCE		1. .862
		/organism="Zea mays"
		/mol_type="mRNA"
		/cultivar="W22"
		/db_xref="taxon:4537"
		/cbase_type="EndospERM of 7-23DAP"
		/clone_id="EndospERM 5"
		/notes="Vector: pBluescript SK-, Site_1: EcoRI, Site_2: XhoI"
BASE COUNT	172 a 216 c 218 g 256 t	
ORIGIN		
Query Match	17.1% Score 672.4; DB 14; Length 862;	
Best Local Similarity	87.4%; Pred. No. 1.4e-139;	
Matches 736; Conservative	0; Mismatches 106; Indels 0; Gaps 0;	
Db	2586 CCTCTGATCTTTTCAGATCGTCTTTCACACAGGTGCTTCGTGGGCTCTTGGTCTTATGAG	2645
	11 CCTCTCAATCTTTTCGATCGTCTTTCACACAGGTCTTCTTGGGCTCTTGGTCAATTGAA	70
Qy	2646 ATCTTCTTCAGCATTCATTCGCTCTTGGTATGGGATAGTGGCGGTCTGAAATTTTG	2705
Db	71 ATTTCTTCAGCAACCACTGCCCTCTCGATGATGGATAGTGTGATGATCTTAAAGTTCGTG	130
Qy	2706 GAAAGATTTTCTTCATCACTCCATCGTATCTTGGACATCATTTCCCTCTGGCT	2765
Db	131 GAAAGTTTTCATCATTTACTTCATGTATACCTTGGACATCTATCCGCTCTTGGC	190
Qy	2766 TACTGTACATTCGCTGCATCTGTATATGACAGGAAATTTATCTCCAGAGCTGAAT	2825
Db	191 TATTGACATTCGCTGCATCTGTGTGTGACAGGAAATTTATCAGCCAGAGCTTAC	250
Qy	2826 AATGTTCGACCTCTGTGTATATGTCATTTTATCTGACATTTTTCATGACAGCATCTTA	2885
Db	251 AATGTTCGACCTCTGTGTATATGTCATTTTATCTGACATTTTTCATGACAGCATCTTG	310
Qy	2886 GAAATGAGATGAGAGTGTGGAATGATGATGATGATGAGAGATGAGAGATGTCGGTC	2945
Db	311 GAAATGAGATGAGAGTGTGATGATGATGATGATGATGAGAGATGAGAGATGTCGGTC	370
Qy	2946 AATGAGAGTGTCTTCACACCTCTTGTGTGTCAGGAGCTTCTCAAGTCTATAGCT	3005
Db	371 AATGAGAGCTGTCTTCACATCTCTTGTGTGTCAGGAGCTTCTCAAGTCTATAGCT	430
Qy	3006 GGTGTGATACAGCTTCACCGTACATCAAAAGGTGAGATGATGAGAGATGTCGACAG	3065
Db	431 GGTGTGATACAGCTTCACCGTACATCAAAAGGTGAGATGATGAGAGATGTCGACAG	490
Qy	3066 CTATATACATTCAAATGAGTACCTTATGATACCTCTTACACCTTCTTATTTAGAC	3125
Db	491 CTATATACATTCAAATGAGTACCTTCTGATACCTCTTACACCTTCTTATTTAGAC	550
Qy	3126 TTCAATGATGTGCTGCTGCGCTTTCAAAATGATCAATTAACGATATGATGATGAGGC	3185
Db	551 TTCAATGATGTGATGCTGCTGCTTCCAAATGATCAATTAACGATATGATGAGGC	610
Qy	3186 CCCCTCTTGGGAGATCTTCTTGGATTTTGGGATGTCATCTTAACTTAACTTAACTTAACT	3245
Db	611 CCCCTCTTGGGAGATCTTCTTGGATTTTGGGATGTCATCTTAACTTAACTTAACTTAACT	670
Qy	3246 AAAGGTTTGTGGAAGCAAAACAGACACCAAGATGTCATGTCGTGTCATTTG	3305
Db	671 AAAGGTCGTGTTGGAGAGCAAGACAGACCAAGATGTCATGTCGTGTCATTTG	730
Qy	3306 CTGGCTTCAATCTTCTGCTCTTGGGTTGGATGATTCCTTCTTGGCAAGATGAT	3365
Db	731 CTGGCTTCAATCTTCTGCTCTTGGGTTGGGATGTCATCTTAACTTAACTTAACTTAACT	790
Qy	3366 GGTGCGCTTCTGGAAGAGTGTGATTTGATTTGACATGATGTCAGTCATGAGCTCC	3425
Db	791 GGTGCGCTTCTGGAAGAGTGTGATTTGATTTGACATGATGTCAGTCATGAGCTCC	850



QY 3426 CC 3427  
 Db 851 CC 852

RESULT 15  
 CC365622/c

LOCUS CC365622 900 bp DNA linear GSS 16-MAY-2003  
 DEFINITION PUHBR527D ZM 0.6 1.0 KB Zea mays genomic clone ZMBMT414J08,  
 genomic survey sequence.

ACCESSION CC365622  
 VERSION CC365622.1 GI:30835022  
 KEYWORDS GSS.

SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Stereophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 900)  
 Whiteley, C.A., Quackenbush, J., Van Aken, S., Uteback, T., Resnick,  
 A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.  
 Maize Genomics Consortium  
 Unpublished  
 CONTACT: Cathy Whiteley  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whiteley@tigr.org  
 Seq primer: TF  
 Clases: sheared ends.

FEATURES  
 source  
 1..900  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone\_1ib="ZMBMT414J08"  
 /clone\_1ib="ZM 0.6 1.0 KB"  
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
 COT selected genomic DNA library"

BASE COUNT 270 a 213 c 198 g 219 t

ORIGIN

Query Match 17.0%; Score 670.6; DB 29; Length 900;  
 Best Local Similarity 88.1%; Pred No. 3.6e-139;  
 Matches 808; Conservative 0; Mismatches 79; Indels 30; Gaps 6;

QY 2977 TGTTCAGGAGCTTCTCAAGTCATAGCTGTGTGATCAAGCTTCACCGTACATCAA 3036  
 |||||  
 Db 900 TGTTCAGGAGCTTCTCAAGTCATAGCTGTGTGATCAAGCTTCACCGTACATCAA 841

QY 3037 AGGGTGGAGATGATGAGAGCTTCTCAGAGCTATATACATTCAAAATGAGACTTATTGA 3096  
 |||||  
 Db 840 AGGGTGGAGATGATGAGAGCTTCTCAGAGCTATATACATTCAAAATGAGACTTATTGA 781

QY 3097 TACCTCCTACGACCTTGTCTTATTGAATTCATGAGTGTGGTGGCGTGGCGTTCAAAATG 3156  
 |||||  
 Db 780 TACCTCCTACGACCTTGTCTTATTGAATTCATGAGTGTGGTGGCGTGGCGTTCAAAATG 721

QY 3157 CGATCAATTAACGATATGAGTATGAGGAGGCGCCCTCTT-TGGGAAGCTATTCTTTGCAATT 3215  
 |||||  
 Db 720 CAATCAATTAATGATATGAGTATGAGGAGGCGCCCTCTTGGGGAAGCTCTTTTGCAATT 661

QY 3216 TGGGTGATTTGTCATCTTTATCCCTTTTCGAAAGTTTGGTTGGAAGGCAAAACAGACA 3275  
 |||||  
 Db 660 TGGGTGATTTGTCATCTTTATCCCTTTTCGAAAGTTTGGTTGGAAGGCAAAACAGACA 601

QY 3276 CCAAGATTGTCATGTCATGTCATTTGCTGGCTGCAATCTTCTCGCTCTTTGGGTT 3335  
 |||||  
 Db 600 CCAAGATTGTCATGTCATGTCATTTGCTGGCTGCAATCTTCTCGCTCTTTGGGTT 541

QY 3336 CGAATTGATCTTCTTGGCAAGATGATGTCGCTTCTTGAGAGATGTGTTGAT 3395

Db 540 CGAATTGACCCCTTCTTGGCAAGATGATGTCGCTTCTTGAGAGATGTGTTGAT 481  
 |||||

QY 3396 TGCACCTA--GGATGTCATGATCAGCTCCGCCAATCGCATGCTTGAAGTATATT 3453  
 |||||

Db 480 TGCACCTAGGGATATCAGACATCAGCTCCGCCAATCAGGTATGATGAATTTT 421  
 |||||

QY 3454 TCTGTGTTTGTCCCATATTCAGT-----TCTGTATTAAGACATGAATGTC 3506  
 |||||

Db 420 GCCGCGTTTGTGCCCATATTTACAGTCCCTCTGTATGATGAAGCAAGAAATGCT 361  
 |||||

QY 3507 CAAGTTTCTTTGATTCATGCTGAACCTTATTATCTGAGATATACCTG-GGAAA 3565  
 |||||

Db 360 CCCGTTCTTTGATTCATGCTGAACCTTATTATCTGAGATATACCTGAGGAAA 301  
 |||||

QY 3566 ATGAGGCTGCGGCAATCCTTGTGAGTTGGGCGGTGAATACAGATATGCAATGTT 3625  
 |||||

Db 300 ATGAGGCTGCGGCAATCCTTGTGAGATGGGCGGTGAATACAGATATGCAATGTT 241  
 |||||

QY 3626 GATTGTGAGCATTTCTTTATTTATTTGTCGCAATATAGATGGCTGAGCCGAAACAG 3685  
 |||||

Db 240 GATTGTGAGCATTTCTTTATTTATTTGTCGCAATATAGATGGCTGAGCCGAAACAG 181  
 |||||

QY 3746 --TGATCTGCGAGGAAACAGAGCACTTCACTTATATGATGCTGCTTATG 3803  
 |||||

Db 120 TGTGATCTGCGAGGAAACAGAGCACTTCACTTATATGATGCTGCTTATG 75  
 |||||

QY 3804 AGGCTGTCTATCATGTCGTCTATATAGAAACAGAAATATAGCTATATCTA 3863  
 |||||

Db 74 ---CAACTCATTAATGATGCTATATATAGAAACAGAAATATGTTGACCAATCTA 18  
 |||||

QY 3864 TAGTTAATTAAGTATG 3880  
 |||||

Db 17 TACTGATTAAGTATG 1

Search completed: December 16, 2003, 00:49:16  
 Job time : 7651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 14:57:33 ; Search time 958 Seconds

(without alignments)  
11090.802 Million cell updates/sec

Title: US-09-720-383C-9

Perfect score: 3936

Sequence: 1 ctctccctcgctcgctcg...ctgagctcgcttgcgaaa 3936

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3936	100.0	3936	21	AAZ58266
2	3936	100.0	3969	21	AAZ99503
3	3936	100.0	3969	21	AAZ99518
4	2592.6	65.9	3776	21	AAZ58263
5	2546.8	64.7	3568	21	AAZ99491
6	2241	56.9	3786	21	AAZ58265
7	2241	56.9	3813	21	AAZ99509
8	2241	56.9	3813	21	AAZ99524

9	1509	38.3	3255	24	AB214725	Arabidopsis thalia
10	1489.8	37.9	3444	22	AAZ58266	Arabidopsis thalia
11	1431.8	36.4	3255	24	AB212754	Arabidopsis thalia
12	1429.4	36.3	3828	19	AAV06566	Arabidopsis thalia
13	1411.4	35.9	3851	19	AAZ58266	Arabidopsis thalia
14	1378.4	35.0	3725	21	AAZ99506	Arabidopsis thalia
15	1378.4	35.0	3725	21	AAZ99506	Arabidopsis thalia
16	1378.4	35.0	3725	21	AAZ99521	Arabidopsis thalia
17	1364	34.7	3753	21	AAZ99515	Arabidopsis thalia
18	1364	34.7	3753	21	AAZ99530	Arabidopsis thalia
19	1364	34.7	3780	21	AAZ99497	Arabidopsis thalia
20	1331.8	33.8	2890	21	AAZ58270	Arabidopsis thalia
21	1327.6	33.7	3673	19	AAV06566	Arabidopsis thalia
22	1326.6	33.7	3603	19	AAV06566	Arabidopsis thalia
23	1321.4	33.6	3704	21	AAZ99533	Arabidopsis thalia
24	1289.4	32.8	3746	21	AAZ99512	Arabidopsis thalia
25	1289.4	32.8	3746	21	AAZ99512	Arabidopsis thalia
26	1289.4	32.8	3773	21	AAZ99494	Arabidopsis thalia
27	1279.4	32.5	3799	24	AAZ58271	Arabidopsis thalia
28	1279.4	32.3	3198	21	AAZ99550	Arabidopsis thalia
29	1269.2	32.2	3614	19	AAV06567	Arabidopsis thalia
30	1254.2	31.9	3517	21	AAZ58268	Arabidopsis thalia
31	1246.2	31.7	3311	19	AAV08373	Arabidopsis thalia
32	1170	29.7	2830	24	AAZ58271	Arabidopsis thalia
33	1137.6	28.9	3207	19	AAV08372	Arabidopsis thalia
34	1136.4	28.9	3328	19	AAV34432	Arabidopsis thalia
35	1118.8	28.4	3747	21	AAZ58271	Arabidopsis thalia
36	1106.2	28.1	3232	22	AAZ58271	Arabidopsis thalia
37	998.8	25.4	2306	19	AAV34433	Arabidopsis thalia
38	884.6	22.5	2248	19	AAV06562	Arabidopsis thalia
39	863	21.9	2033	19	AAV08381	Arabidopsis thalia
40	859.6	21.8	2125	21	AAZ58269	Arabidopsis thalia
41	856.4	21.8	2055	21	AAZ58269	Arabidopsis thalia
42	749.2	19.0	1860	21	AAZ58271	Arabidopsis thalia
43	741.4	18.8	1029	21	AAZ58271	Arabidopsis thalia
44	701.8	17.8	1734	21	AAZ58271	Arabidopsis thalia
45	664.2	16.9	1741	19	AAV06569	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
ID AAZ58266 standard; cDNA; 3936 BP.  
AC AAZ58266;  
XX  
DT 08-MAY-2000 (first entry)  
XX  
DE Corn cellulose synthase cDNA clone p0122.ckmh70rc.  
XX  
XX Corn; maize; cellulose synthase; transgenic plant; ss.  
XX  
XX Zea mays.  
OS  
FH Key Location/Qualifiers  
FT CDS 144..3404  
FT /tag= a  
XX  
XX  
XX WO200004166-A2.  
XX  
XX 27-JAN-2000.  
XX  
XX 13-JUL-1999; 99WO-US15871.  
XX  
XX 14-JUL-1998; 98US-0092844.  
XX  
XX (DUPLO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Allen SM, Fader GW, Falco SC, Kinney AJ, Lightner JB, Miao G,  
XX Rafael JA, Thorpe CJ;



QY 1741 ATAAAGAAAGCTGTGCTATGAAATGCAATGGTCCGAGTCTCTGCTGTAATAACAAATGCTC 1800  
DB 1741 ATAGAAAGCTGTGCTATGAAATGCAATGGTCCGAGTCTCTGCTGTAATAACAAATGCTC 1800  
QY 1801 CATATTTGTTAACTTGGATGTTGATCACTACATCAACACAGAGGCTATTAAGAGAG 1860  
DB 1801 CATATTTGTTAACTTGGATGTTGATCACTACATCAACACAGAGGCTATTAAGAGAG 1860  
QY 1861 CAATGTGTTTATGATGGAACCTTTAATGAGAAAGAGTTTGTATGTACAGTTCCCTC 1920  
DB 1861 CAATGTGTTTATGATGGAACCTTTAATGAGAAAGAGTTTGTATGTACAGTTCCCTC 1920  
QY 1921 AAAAGTTGATGGAGTTGATGCGCAATGCGATATGCTAACCGGAATGTGCTTTTGG 1980  
DB 1921 AAAAGTTGATGGAGTTGATGCGCAATGCGATATGCTAACCGGAATGTGCTTTTGG 1980  
QY 1981 ATATCAACATGAAAGGTTTGGATGGATTCAGGGTCCAAATTTATGTGTTACATGATG 2040  
DB 1981 ATATCAACATGAAAGGTTTGGATGGATTCAGGGTCCAAATTTATGTGTTACATGATG 2040  
QY 2041 TATTTAGAGGAGGCAATTAATGTTATGATGCCCCCAAAACAAAGAACCAATCAA 2100  
DB 2041 TATTTAGAGGAGGCAATTAATGTTATGATGCCCCCAAAACAAAGAACCAATCAA 2100  
QY 2101 GSACTTGAACCTGTGCGCCAAAGTGTGCTTTGCTGCTGCTTGGCAATAGAGAGC 2160  
DB 2101 GSACTTGAACCTGTGCGCCAAAGTGTGCTTTGCTGCTGCTTGGCAATAGAGAGC 2160  
QY 2161 AAAAGAGACTACCAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
DB 2161 AAAAGAGACTACCAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
QY 2221 AGAACCAATCCCTGCTCATATGCTCTTGTGTAATTTGACGAGCTGCTCCAGAGCTGAGA 2280  
DB 2221 AGAACCAATCCCTGCTCATATGCTCTTGTGTAATTTGACGAGCTGCTCCAGAGCTGAGA 2280  
QY 2281 ATGAAGAAGCGGCTTGTGAATTCACAAATTTGAAAGAAATTTGGCCCAATCTTCTG 2340  
DB 2281 ATGAAGAAGCGGCTTGTGAATTCACAAATTTGAAAGAAATTTGGCCCAATCTTCTG 2340  
QY 2341 TTTTGTATACATCCACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAGTCTGCTT 2400  
DB 2341 TTTTGTATACATCCACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAGTCTGCTT 2400  
QY 2401 CTCTTTGAAAGAGCTATACATGTCATTAATGTTGTTATGAAACAGACAGCTGGG 2460  
DB 2401 CTCTTTGAAAGAGCTATACATGTCATTAATGTTGTTATGAAACAGACAGCTGGG 2460  
QY 2461 GAAAAGAGATGGCTGATCTATGATCAGTTACAGAGATATTTCTAATGCTGTTCAGAA 2520  
DB 2461 GAAAAGAGATGGCTGATCTATGATCAGTTACAGAGATATTTCTAATGCTGTTCAGAA 2520  
QY 2521 TGCAATGTCATGCTGAGGAGTCAATTTTACTGCAATCCTTAAACGGGTTGCAATCAAGGTT 2580  
DB 2521 TGCAATGTCATGCTGAGGAGTCAATTTTACTGCAATCCTTAAACGGGTTGCAATCAAGGTT 2580  
QY 2581 CTGCACTCTGATCTTTCAATGCTTTCACAGAGTCTTCTGAGGCTCTTGGGTCTA 2640  
DB 2581 CTGCACTCTGATCTTTCAATGCTTTCACAGAGTCTTCTGAGGCTCTTGGGTCTA 2640  
QY 2641 TTGAGATCTTCTTCAAGCAATCTTGCCTTTGGTATGGGTATGGTGGCGGTCTGGAAT 2700  
DB 2641 TTGAGATCTTCTTCAAGCAATCTTGCCTTTGGTATGGGTATGGTGGCGGTCTGGAAT 2700  
QY 2701 TTTTGAAGAGATTTCTACATCACTCATGATGTTTATGAGCAATCTATTTCCCTCT 2760  
DB 2701 TTTTGAAGAGATTTCTACATCACTCATGATGTTTATGAGCAATCTATTTCCCTCT 2760  
QY 2761 TGGCTTACTGTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
DB 2761 TGGCTTACTGTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
QY 2821 TGAATTAATGTTGCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880

DB 2821 TGAATTAATGTTGCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
QY 2881 TCTAGAAATGAGATGAGATGAGTGTGGAATTTGATGATGCTGTGAGAGATGAGAGTCT 2940  
DB 2881 TCTAGAAATGAGATGAGATGAGTGTGGAATTTGATGATGCTGTGAGAGATGAGAGTCT 2940  
QY 2941 GGGTCAATTTGAGAGTGTGCTTCAACCTCTTGTGCTGTGCTTCAAGAGCTTCTCAAGGCTCA 3000  
DB 2941 GGGTCAATTTGAGAGTGTGCTTCAACCTCTTGTGCTGTGCTTCAAGAGCTTCTCAAGGCTCA 3000  
QY 3001 TAGCTGTGTTGATACAAAGCTTCAACCGTACATCAAGAGGAGGAGATGATGAGAGTCT 3060  
DB 3001 TAGCTGTGTTGATACAAAGCTTCAACCGTACATCAAGAGGAGGAGATGATGAGAGTCT 3060  
QY 3061 CAGAGCTATATACATTCATATGAGATCTTATTTGATACCTTACCACTTGTCTTAT 3120  
DB 3061 CAGAGCTATATACATTCATATGAGATCTTATTTGATACCTTACCACTTGTCTTAT 3120  
QY 3121 TGAATCTTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
DB 3121 TGAATCTTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
QY 3181 GGGGCCCCCTCTTGGGAGAGCTATCTTGTGATTTGGGATGTTGCCATCTTATCCCT 3240  
DB 3181 GGGGCCCCCTCTTGGGAGAGCTATCTTGTGATTTGGGATGTTGCCATCTTATCCCT 3240  
QY 3241 TTCTCAAGGTTTGGTTGGAAGGCAAAACAGAGACCAACGATGTCATCTGTGCTCA 3300  
DB 3241 TTCTCAAGGTTTGGTTGGAAGGCAAAACAGAGACCAACGATGTCATCTGTGCTCA 3300  
QY 3301 TTCTGCTGCTTCAATCTTCTGCTCCTTGGGTTGGATGATCTTCTTGGAGAG 3360  
DB 3301 TTCTGCTGCTTCAATCTTCTGCTCCTTGGGTTGGATGATCTTCTTGGAGAG 3360  
QY 3361 ATGATGCTGCTTCTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
DB 3361 ATGATGCTGCTTCTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
QY 3421 CTCCCCCAATCTGATATGCTTGAAGTATTTTCTGCTGCTTCTCCCATATTCAGTGT 3480  
DB 3421 CTCCCCCAATCTGATATGCTTGAAGTATTTTCTGCTGCTTCTCCCATATTCAGTGT 3480  
QY 3481 CTGTATGATGAGATCAATGAATGTCCTCAAGTTCTTTTGAATCAATGATGTAACCTTAA 3540  
DB 3481 CTGTATGATGAGATCAATGAATGTCCTCAAGTTCTTTTGAATCAATGATGTAACCTTAA 3540  
QY 3541 TATCTGAGATATATCTGGGGGAAATGAGAGCTGCGGCAATCTTGTGCAAGTTGGGCG 3600  
DB 3541 TATCTGAGATATATCTGGGGGAAATGAGAGCTGCGGCAATCTTGTGCAAGTTGGGCG 3600  
QY 3601 TGAATACAGATATGCAATGTTGATGTTGAGATGTTGAGATGTTGAGATGTTGAGATGTT 3660  
DB 3601 TGAATACAGATATGCAATGTTGATGTTGAGATGTTGAGATGTTGAGATGTTGAGATGTT 3660  
QY 3661 TAGATGGCTGAGCGAAACAGAGATTTTGAATTTGAGATGTTGAGATGTTGAGATGTT 3720  
DB 3661 TAGATGGCTGAGCGAAACAGAGATTTTGAATTTGAGATGTTGAGATGTTGAGATGTT 3720  
QY 3721 TGGTTCCTAATAAGGAGGAGGAGATGATCTGCAAGTGAACAGAGACCACTGCACTT 3780  
DB 3721 TGGTTCCTAATAAGGAGGAGGAGATGATCTGCAAGTGAACAGAGACCACTGCACTT 3780  
QY 3781 ATTTATGATGCTGTTCAATGAGAGGCTTGTCAATTAATGATGATGATGATGATGATGAT 3840  
DB 3781 ATTTATGATGCTGTTCAATGAGAGGCTTGTCAATTAATGATGATGATGATGATGATGAT 3840  
QY 3841 AACAGAAATATGAGATTAATCTATATGATTAATTAAGATGATTAATGATGATGATGAT 3900  
DB 3841 AACAGAAATATGAGATTAATCTATATGATTAATTAAGATGATTAATGATGATGATGAT 3900  
QY 3901 TGTGTACTGTATCATCTGAGTGTGTTTGTGAAAA 3936

Query Match	Best Local Similarity	Score	DB	Length	3969;
Match 3336; Conservative	100.0%;	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
				Gaps 0;	

QY	1	CTTCTCCCTCGTCCGATGCGAGCTGGCCGGCTTCGGGTGAGAAACCACTCGGAGG	60
Db	1	CTTCTCCCTCGTCCGATGCGAGCTGGCCGGCTTCGGGTGAGAAACCACTCGGAGG	60
QY	61	ATGAGGATCTGCTCTAGAGTGAAGGAGCTACGGTACAGTATCCCTTCGCTCGTCCGCG	120
Db	61	ATGAGGATCTGCTCTAGAGTGAAGGAGCTACGGTACAGTATCCCTTCGCTCGTCCGCG	120
QY	121	GCGGAGTGAAGGGGAGGAAACGATGAGGCGAGCGCCGGCTGTGGCCGGCTCCACA	180
Db	121	GCGGAGTGAAGGGGAGGAAACGATGAGGCGAGCGCCGGCTGTGGCCGGCTCCACA	180
QY	181	ACCGCAACGAGCTGTCGTATCCGCGCGCACGGCGATCCCGGGCCGAAGCCGCGCGG	240
Db	181	ACCGCAACGAGCTGTCGTATCCGCGCGCACGGCGATCCCGGGCCGAAGCCGCGCGG	240
QY	241	AGCAGAACGGGCAAGTGTGCCAATTGGCGGCGACGATCCGCGCTTGCCCGCGGGGG	300
Db	241	AGCAGAACGGGCAAGTGTGCCAATTGGCGGCGACGATCCGCGCTTGCCCGCGGGGG	300
QY	301	ACCCCTTCGTGGCTGCAACGAGTGGCCCTTCCCGCTGGCGGACCTGTAACGAATACG	360
Db	301	ACCCCTTCGTGGCTGCAACGAGTGGCCCTTCCCGCTGGCGGACCTGTAACGAATACG	360
QY	361	AGCGCCGGGAGGGACGCGAATCTGCCCGCAGTCCAAACTCGATACAAAGCCCTCAAG	420
Db	361	AGCGCCGGGAGGGACGCGAATCTGCCCGCAGTCCAAACTCGATACAAAGCCCTCAAG	420
QY	421	GCTGCGCAACGTGTACCCGGTGAACGAGAGGAGACGGCGTGTGATGACTTGACACGAGT	480
Db	421	GCTGCGCAACGTGTACCCGGTGAACGAGAGGAGACGGCGTGTGATGACTTGACACGAGT	480
QY	481	TCAACTGGGAGCGGCAATGACTCGAGTCTGTGGCGGAGTCCATCTCTACGCGCACATGA	540
Db	481	TCAACTGGGAGCGGCAATGACTCGAGTCTGTGGCGGAGTCCATCTCTACGCGCACATGA	540
QY	541	GCTACGCGCGGTGAGGTGACCTTAATGGCGCGCACAACTTTCCACTCAACCCCAATG	600
Db	541	GCTACGCGCGGTGAGGTGACCTTAATGGCGCGCACAACTTTCCACTCAACCCCAATG	600
QY	601	TTTCCATCTCTACCAACGGGCAATTGGTGTATGACATCCACCCGAGACGACGCGCTGG	660
Db	601	TTTCCATCTCTACCAACGGGCAATTGGTGTATGACATCCACCCGAGACGACGCGCTGG	660
QY	661	TGCGCTTCCTTATGGGTGGTGGGGGAAAGGATTAATCCCTTCCTTATGGGATCCCA	720
Db	661	TGCGCTTCCTTATGGGTGGTGGGGGAAAGGATTAATCCCTTCCTTATGGGATCCCA	720
QY	721	GCTTACCTGTGCAACCCAGGCTCTATGCAACCCATCCAGGATCTTGCTGCATATGGTATG	780
Db	721	GCTTACCTGTGCAACCCAGGCTCTATGCAACCCATCCAGGATCTTGCTGCATATGGTATG	780
QY	781	GTAGTGTGCTTGGAGAGGATCGATTGGAGAAATTGGAAACGAGACAAAGAGATGCACC	840
Db	781	GTAGTGTGCTTGGAGAGGATCGATTGGAGAAATTGGAAACGAGACAAAGAGATGCACC	840
QY	841	AGAAGGGGAAATGATGGTGGTGTATATGTTGAGAGATGCTGATCTACATTAATGATG	900
Db	841	AGAAGGGGAAATGATGGTGGTGTATATGTTGAGAGATGCTGATCTACATTAATGATG	900
QY	901	AAGCAAGACAACAATGTCCAGAGAAATTCACATCTTCATCAAGCCAGATTATCATATA	960
Db	901	AAGCAAGACAACAATGTCCAGAGAAATTCACATCTTCATCAAGCCAGATTATCATATA	960
QY	961	GGATGATTAATCATTAATTCGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCT	1020
Db	961	GGATGATTAATCATTAATTCGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCT	1020
QY	1021	ATCCGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1080
Db	1021	ATCCGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1080
QY	1081	CCATGCTTGGATTTCTGATCAATTCCCAAGTGGTTCCTTATGAGAGAGACCTTACC	1140

Db	1081	CCAATGCTTGGATTCTTGAATCGAATTCCAAAGTGGTTCCTAATTGAGAGAGACTTAACC	1140
Qy	1141	TAGACCGGCTGTCACTGAGTTTCCAGCAAGAAAGCCAGCAATCTCGAATTG	1200
Db	1141	TAGACCGGCTGTCACTGAGTTTCCAGCAAGAAAGCCAGCAATCTCTCAATTG	1200
Qy	1201	ATTTCTTTGTGAGTACGGTGTGATCCCTTAAAGAAACCTCTTGGTCACAAAATCTG	1260
Db	1201	ATTTCTTTGTGAGTACGGTGTGATCCCTTAAAGAAACCTCTTGGTCACAAAATCTG	1260
Qy	1261	TTCTATCATCTCTTCCGTGGAATTATCTGTGTAATAAGTTTCTTCTGTAATGTTCTGATG	1320
Db	1261	TTCTATCATCTCTTCCGTGGAATTATCTGTGTAATAAGTTTCTTCTGTAATGTTCTGATG	1320
Qy	1321	ATGTGCTGCAATGCTAAAGTTTGAAAGCAATTATCTGAAACATCTGAATTTGCCAAAGAAAT	1380
Db	1321	ATGTGCTGCAATGCTAAAGTTTGAAAGCAATTATCTGAAACATCTGAATTTGCCAAAGAAAT	1380
Qy	1381	GGGTTCCTTTCTGCAAAAGGTATCAATTGTAACCTCCGGCTCCAGAGTGTACTTCCAAAC	1440
Db	1381	GGGTTCCTTTCTGCAAAAGGTATCAATTGTAACCTCCGGCTCCAGAGTGTACTTCCAAAC	1440
Qy	1441	AGAAAGTATGACTACTTGAAGAACAAGTGGACGACAACTTTGTTTAGGAGAGAGAGACAA	1500
Db	1441	AGAAAGTATGACTACTTGAAGAACAAGTGGACGACAACTTTGTTTAGGAGAGAGAGACAA	1500
Qy	1501	TGAAGAGAGATAGAGAAATTCAGAGTGAATCAATGCTTAGTTGCCAAAGCCGACA	1560
Db	1501	TGAAGAGAGATAGAGAAATTCAGAGTGAATCAATGCTTAGTTGCCAAAGCCGACA	1560
Qy	1561	AAAGTTCTGAAAGAGATGACAAATGCAAGATGCAACCCCTGGCTGGAACAATGTTTC	1620
Db	1561	AAAGTTCTGAAAGAGATGACAAATGCAAGATGCAACCCCTGGCTGGAACAATGTTTC	1620
Qy	1621	GTGATCATCTCGAATGATGATCAGGCTTCCTTGGCAAAAGGAGAGCCCTTACATGTGAGG	1680
Db	1621	GTGATCATCTCGAATGATGATCAGGCTTCCTTGGCAAAAGGAGAGCCCTTACATGTGAGG	1680
Qy	1681	GAAATGAACTGCAACGATTTGATTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC	1740
Db	1681	GAAATGAACTGCAACGATTTGATTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC	1740
Qy	1741	ATAAGAAAGCTGGTGTATGAAATGATGTCGAGTCTCTGCTGTACTTAACAATGCTC	1800
Db	1741	ATAAGAAAGCTGGTGTATGAAATGATGTCGAGTCTCTGCTGTACTTAACAATGCTC	1800
Qy	1801	CATATTTGTAACTTGGATTTGTATCTACTATCAATAACAAGCAAGGCTATAAAGAAAG	1860
Db	1801	CATATTTGTAACTTGGATTTGTATCTACTATCAATAACAAGCAAGGCTATAAAGAAAG	1860
Qy	1861	CAATGTGTTTATGATGAGCCCTTATCAGAGAAAGAAAGTGTGCTATGTACAGTTCCCTC	1920
Db	1861	CAATGTGTTTATGATGAGCCCTTATCAGAGAAAGAAAGTGTGCTATGTACAGTTCCCTC	1920
Qy	1921	AAAAGTTTGAATGGATTTGATGCGCCATGACCGATATGCTAACCGAATGTTGCTTTTGG	1980
Db	1921	AAAAGTTTGAATGGATTTGATGCGCCATGACCGATATGCTAACCGAATGTTGCTTTTGG	1980
Qy	1981	ATATCAACATGAAAGTTTGGATGATTTCAAGGTCCTCAATTTATGTTGTACTGGATGTG	2040
Db	1981	ATATCAACATGAAAGTTTGGATGATTTCAAGGTCCTCAATTTATGTTGTACTGGATGTG	2040
Qy	2041	TATTTAGAAAGGAGCATTTATTTGTTTATGATGCCCCCAAAACAAGAGCAACATGAA	2100
Db	2041	TATTTAGAAAGGAGCATTTATTTGTTTATGATGCCCCCAAAACAAGAGCAACATGAA	2100
Qy	2101	GGAATTTGCAACTGTGAGCCCAAGTGTGCTTTGCTGTGCTTTGGCAATAGAAAGC	2160
Db	2101	GGAATTTGCAACTGTGAGCCCAAGTGTGCTTTGCTGTGCTTTGGCAATAGAAAGC	2160
Qy	2161	AAAAAGAGCTACCAAAACCAAAACAGAGAGAAAGTTATTTTCAAGAAAGAG	2220
Db	2161	AAAAAGAGCTACCAAAACCAAAACAGAGAGAAAGTTATTTTCAAGAAAGAG	2220

Dp	2161	AAAAGAGACTACCAAAACCCAAAACAGAAAGAAAAAGTATATATTTTTCAGAAAGAG	2220
Qy	2221	AGAAACAATCCCTGCATATGCTCTTGTAATTTGACGAGCTGCTCCAGAGCTGAGA	2280
Dp	2221	AGAAACAATCCCTGCATATGCTCTTGTAATTTGACGAGCTGCTCCAGAGCTGAGA	2280
Qy	2281	ATGAAAAGGCGGTATTTGTAAATTCACAAAAATTAGAAAAGAAATTTGGCCAAATCTTCTG	2340
Dp	2281	ATGAAAAGGCGGTATTTGTAAATTCACAAAAATTAGAAAAGAAATTTGGCCAAATCTTCTG	2340
Qy	2341	TTTTTGTATACATCACACTTCTCGAAGATGATGGAACCTTGAAAGATGCAAGTCCCTGCT	2400
Dp	2341	TTTTTGTATACATCACACTTCTCGAAGATGATGGAACCTTGAAAGATGCAAGTCCCTGCT	2400
Qy	2401	CTCTTTGAAAAGACTATACATGTCATTAGTTGTGTTATGAGACAAGAACAAGCTGGG	2460
Dp	2401	CTCTTTGAAAAGACTATACATGTCATTAGTTGTGTTATGAGACAAGAACAAGCTGGG	2460
Qy	2461	GAAAAAGATTGGCTGTGATCTATATGATTCAGTTACAGAAAGATTTCTAACTGTGTTCAAGA	2520
Dp	2461	GAAAAAGATTGGCTGTGATCTATATGATTCAGTTACAGAAAGATTTCTAACTGTGTTCAAGA	2520
Qy	2521	TGCATTGTCAATGGTGGCGGTCAATTTTATGCAATCTTAAAGGGTGTGCAATTCAAAGGTT	2580
Dp	2521	TGCATTGTCAATGGTGGCGGTCAATTTTATGCAATCTTAAAGGGTGTGCAATTCAAAGGTT	2580
Qy	2581	CTGACACTCTGAATCTTTTCAGATCGTCTTACACAGGTGCTTGGTGGGCTCTTGGGTCTA	2640
Dp	2581	CTGACACTCTGAATCTTTTCAGATCGTCTTACACAGGTGCTTGGTGGGCTCTTGGGTCTA	2640
Qy	2641	TTGAGATCTTCTTCAGCAATCATATGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAT	2700
Dp	2641	TTGAGATCTTCTTCAGCAATCATATGCCCTCTTTGGTATGGGTATGGTGGCGGTCTGAAT	2700
Qy	2701	TTTTGAAAAGATTTTCTCAATCAATCCATGATGATCTTGGACATCTAATCCCTCT	2760
Dp	2701	TTTTGAAAAGATTTTCTCAATCAATCCATGATGATCTTGGACATCTAATCCCTCT	2760
Qy	2761	TGGCTTACCTGATCATATGCCCTGCATCTGTTTATGACAGGAAATTTATCACTCCAGAGC	2820
Dp	2761	TGGCTTACCTGATCATATGCCCTGCATCTGTTTATGACAGGAAATTTATCACTCCAGAGC	2820
Qy	2821	TGAATTAATGTTGCCAGCTGTGGTTATGTCAATGTCACCTTTTATCTGCAATTTTGGTCAAGACA	2880
Dp	2821	TGAATTAATGTTGGCAGCTGTGGTTATGTCAATGTTTATCTGCAATTTTGGTCAAGACA	2880
Qy	2881	TCCAGAAATGAGATGAGATGAGTGGTGGAAATTGATGACGTGGGAGAAATGAGCAAGTCT	2940
Dp	2881	TCCAGAAATGAGATGAGATGAGTGGTGGAAATTGATGACGTGGGAGAAATGAGCAAGTCT	2940
Qy	2941	GGGTCAATTTGAGAGGTGTCTCTCAACCTCTTTGCTGTGTTCCAGGGAATTTCTCAAGTCA	3000
Dp	2941	GGGTCAATTTGAGAGGTGTCTCTCAACCTCTTTGCTGTGTTCCAGGGAATTTCTCAAGTCA	3000
Qy	3001	TAGCTGTGTTGATPACAGCTTCAACCTGATCAATCAAAAGGTGAGATGATGAGAGTCT	3060
Dp	3001	TAGCTGTGTTGATPACAGCTTCAACCTGATCAATCAAAAGGTGAGATGATGAGAGTCT	3060
Qy	3061	CAGAGCTATPACATTCAAATGGAACCTTATATGATACCTTCAACCTTGTCTTCAAT	3120
Dp	3061	CAGAGCTATPACATTCAAATGGAACCTTATATGATACCTTCAACCTTGTCTTCAAT	3120
Qy	3121	TGAATCTCATTTGATGTGTGCTGCTGAGCGTTTCAAAATGCGATCAATACGATATGAGTCAAT	3180
Dp	3121	TGAATCTCATTTGATGTGTGCTGCTGAGCGTTTCAAAATGCGATCAATACGATATGAGTCAAT	3180
Qy	3181	GGGGCCCCCTCTTTGGAGAGCTATCTTTGCAATTTGGGTATGTGCAATCTTATCCCT	3240
Dp	3181	GGGGCCCCCTCTTTGGAGAGCTATCTTTGCAATTTGGGTATGTGCAATCTTATCCCT	3240
Qy	3241	TTTCCTCAAGGTTGGTTGGAAGGCAAAACAGACACCAACGATGTGCATGCTGTGCTCA	3300
Dp	3241	TTTCCTCAAGGTTGGTTGGAAGGCAAAACAGACACCAACGATGTGCATGCTGTGCTCA	3300





Db 421 GCTGCAACGTGTGACGGGTGACGAGAGAGAGACGGCGTGCATGACCTGGACAACGAGT 480  
Qy 481 TCMACTGGAGACGGCCATGACTGCGAGCTGTGGCCGAGTCCATGTCTTACGGCCACATGA 540  
Db 481 TCMACTGGAGACGGCCATGACTGCGAGCTGTGGCCGAGTCCATGTCTTACGGCCACATGA 540  
Qy 541 GCTACGGCCGTGGAGGTGACCTTAAATGGCGGCCCAAGCTTTCAGTCCAAACCCCAATG 600  
Db 541 GCTACGGCCGTGGAGGTGACCTTAAATGGCGGCCCAAGCTTTCAGTCCAAACCCCAATG 600  
Qy 601 TTCACCTCTCAACCAACGGGCAAAATGTGTGATGACATCCACCGAGACAGACGGCTGG 660  
Db 601 TTCACCTCTCAACCAACGGGCAAAATGTGTGATGACATCCACCGAGACAGACGGCTGG 660  
Qy 661 TGCCCTCTTTCATGGGTGGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCCA 720  
Db 661 TGCCCTCTTTCATGGGTGGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCCA 720  
Qy 721 GCTTACCTGTGCAACCCAGGTCTATGACCCATCCAGGATCTTGTGATGAGGTATG 780  
Db 721 GCTTACCTGTGCAACCCAGGTCTATGACCCATCCAGGATCTTGTGATGAGGTATG 780  
Qy 781 GTAGGTCTTGGAGAGAGACGGATGAGAAATTTGGAACAGAGACAAAGAGATGACAC 840  
Db 781 GTAGGTCTTGGAGAGAGACGGATGAGAAATTTGGAACAGAGACAAAGAGATGACAC 840  
Qy 841 AGACGGGGAATGATGT 900  
Db 841 AGACGGGGAATGATGT 900  
Qy 901 AAGCAAGCAACAACTGTCCAGGAAATTTCCATCTTCCATCAAGCCAGATTAACTCATTA 960  
Db 901 AAGCAAGCAACAACTGTCCAGGAAATTTCCATCTTCCATCAAGCCAGATTAACTCATTA 960  
Qy 961 GGATGATATATCATTAATTTGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020  
Db 961 GGATGATATATCATTAATTTGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020  
Qy 1021 ATCCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
Db 1021 ATCCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
Qy 1081 CCAATGTCTTGGATTTTGTATCAATTTCCAAAGTGGTTCCTATTGTGAGAGAGACTTACC 1140  
Db 1081 CCAATGTCTTGGATTTTGTATCAATTTCCAAAGTGGTTCCTATTGTGAGAGAGACTTACC 1140  
Qy 1081 CCAATGTCTTGGATTTTGTATCAATTTCCAAAGTGGTTCCTATTGTGAGAGAGACTTACC 1140  
Db 1081 CCAATGTCTTGGATTTTGTATCAATTTCCAAAGTGGTTCCTATTGTGAGAGAGACTTACC 1140  
Qy 1141 TAGACCGGCTGCACTGAGGTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
Db 1141 TAGACCGGCTGCACTGAGGTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
Qy 1201 ATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
Db 1201 ATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
Qy 1261 TTTTATCTATCTTTTGGTGTGATTAATCTGTGATTAAGGTTTCTGTATGTTTCTGATG 1320  
Db 1261 TTTTATCTATCTTTTGGTGTGATTAATCTGTGATTAAGGTTTCTGTATGTTTCTGATG 1320  
Qy 1321 ATGGGTCTGCAATGCTTAACGTTTGAAGATTAATCTGAACCAATCTGAATTTTGAAGAAAT 1380  
Db 1321 ATGGGTCTGCAATGCTTAACGTTTGAAGATTAATCTGAACCAATCTGAATTTTGAAGAAAT 1380  
Qy 1381 GGGTTCCTTTCGCAACCGGTACATATTGAACCTCGGCTCCAAAGTGGTATCTTCCAAAC 1440  
Db 1381 GGGTTCCTTTCGCAACCGGTACATATTGAACCTCGGCTCCAAAGTGGTATCTTCCAAAC 1440  
Qy 1441 AGAAGATGACTACTTTGAAGACAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGACAA 1500  
Db 1441 AGAAGATGACTACTTTGAAGACAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGACAA 1500  
Qy 1501 TGAAGAGAGATGAGAAATTCAGAGTGAAGAAATGAAGCTTATGTCGCAAGGCCAGA 1560  
Db 1501 TGAAGAGAGATGAGAAATTCAGAGTGAAGAAATGAAGCTTATGTCGCAAGGCCAGA 1560

Qy 1561 AAGTTCCTGAAGAGATGAGCAATGCAAGTGAACCCCTGCGCTGGAAACAAATGTTG 1620  
Db 1561 AAGTTCCTGAAGAGATGAGCAATGCAAGTGAACCCCTGCGCTGGAAACAAATGTTG 1620  
Qy 1621 GTGATCATCTCGAATGATTCAGGTCTTCTGTCGCAAGCGAGGCTTGACTGTGAG 1680  
Db 1621 GTGATCATCTCGAATGATTCAGGTCTTCTGTCGCAAGCGAGGCTTGACTGTGAG 1680  
Qy 1681 GAAATGAACCTGCCAATGGTTTATGTTTCTAGAGAAACGACAGGCTTAACATC 1740  
Db 1681 GAAATGAACCTGCCAATGGTTTATGTTTCTAGAGAAACGACAGGCTTAACATC 1740  
Qy 1741 ATAGAAAGCTGGTGTATGAAATGATTTGGTCCGAGTCTGTGCTATCTAACAATGCTC 1800  
Db 1741 ATAGAAAGCTGGTGTATGAAATGATTTGGTCCGAGTCTGTGCTATCTAACAATGCTC 1800  
Qy 1801 CATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db 1801 CATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Qy 1861 CAATGTGTTTATGATGACCCCTTACTAGAAAGAGTTTGTATGTAACAGTTCCTC 1920  
Db 1861 CAATGTGTTTATGATGACCCCTTACTAGAAAGAGTTTGTATGTAACAGTTCCTC 1920  
Qy 1921 AAAATTTGATGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980  
Db 1921 AAAATTTGATGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980  
Qy 1981 ATATCAACATGAAAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Db 1981 ATATCAACATGAAAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Qy 2041 TATTTAGAGGCAAGCATTAATATGTTATGATGATGATGATGATGATGATGATGATGATG 2100  
Db 2041 TATTTAGAGGCAAGCATTAATATGTTATGATGATGATGATGATGATGATGATGATGATG 2100  
Qy 2101 GGACTTGCACTGCTGGCCCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2160  
Db 2101 GGACTTGCACTGCTGGCCCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2160  
Qy 2161 AAAAAGAGACTACCAAAACCCCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
Db 2161 AAAAAGAGACTACCAAAACCCCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
Qy 2221 AGAACCAATCCCTGCAATATGCTGTGTGAAATTTGACGAAGCTGCTCAAGAGCTGAGA 2280  
Db 2221 AGAACCAATCCCTGCAATATGCTGTGTGAAATTTGACGAAGCTGCTCAAGAGCTGAGA 2280  
Qy 2281 ATGAAAAGCCGGTATTTGTAATCAACAAAATTTAGAAAAGAAATTTGGCCAAATCTTCTG 2340  
Db 2281 ATGAAAAGCCGGTATTTGTAATCAACAAAATTTAGAAAAGAAATTTGGCCAAATCTTCTG 2340  
Qy 2341 TTTTGTGATCATCAACCTTCTCGAAGAAATGATGAAACCTTGAAGAGTCTGCTT 2400  
Db 2341 TTTTGTGATCATCAACCTTCTCGAAGAAATGATGAAACCTTGAAGAGTCTGCTT 2400  
Qy 2401 CTCTTTGAAAGAGACTATACATGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2460  
Db 2401 CTCTTTGAAAGAGACTATACATGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2460  
Qy 2461 GAAAGAGATTTGCTGATCTATGATCAAGTTACAGAAAGATTTCTAACTGTTTCAAGA 2520  
Db 2461 GAAAGAGATTTGCTGATCTATGATCAAGTTACAGAAAGATTTCTAACTGTTTCAAGA 2520  
Qy 2521 TGCAATGTCATGTTGGCGGTCAATTTTACGCAATACCTTAAACGGGTGCAATCAAGTT 2580  
Db 2521 TGCAATGTCATGTTGGCGGTCAATTTTACGCAATACCTTAAACGGGTGCAATCAAGTT 2580  
Qy 2581 CTGCACTCTGAATCTTTCAAGATGCTTCAACAGTGGCTTGGTGGCTTGGTGGCTT 2640  
Db 2581 CTGCACTCTGAATCTTTCAAGATGCTTCAACAGTGGCTTGGTGGCTTGGTGGCTT 2640





Db 2239 GATTAAGAACTTTTAAAGAAAAGAAATCAAGCCCTGCATATGCTCTTGGAAT 2298  
 Qy 2255 TGACGAAGCTGCTCCAGAGCTGAGATGAAAAAGCCGGATTTGTAATCAAAAT 2314  
 Db 2299 TGATGAAGCCGCTCCAGAGCTGAAAAATGAAAAGCTAGATTTGTAATCAAGAA 2358  
 Qy 2315 AGAAAAGAAATTTGGCCAACTCTTCTGTTTTGTTATCATCACTTCGAGATGTTG 2374  
 Db 2359 GAAAAAGAAATTTGGCAGCTCTTCAGTTTTGTTGATCCACACTTCTGAGATGTTG 2418  
 Qy 2375 AACCTTGAAGAGTCAAGTCCGCTCTCTTTGAAAAGCATATCATGTCATGTTG 2434  
 Db 2419 AACCTGAAAGAGTCCAGTCCAGCTTCTCTTGAAGAAAGTATACATGATCAGTTG 2478  
 Qy 2435 TGGTTATGAAGACAAGACAGACTGGGAAAAAGATGAGTGGATCTATGATCAGTTAC 2494  
 Db 2479 TCGATATGAAGACAAGAGCTGGGAAAAAGATTTGGTTGATTTATGATCAGTAC 2538  
 Qy 2495 AGAAGATATTTCACTGCTTCAAGATGCAATGCTGTTGGCCGTCATTTACTGCA 2554  
 Db 2539 AGAAGATATTTCACTGCTTCAAGATGCAATGCTGTTGGCCGTCATTTACTGCA 2598  
 Qy 2555 ACCTAAACGGGTTGATCAAGGTTCTGACCTCGAATCTTTCAGATCGCTTACCA 2614  
 Db 2599 ACCTAAACGGGCGCCCTTCAAGGTTCCGACCTCAATCTTCCGATCGCTTACCA 2658  
 Qy 2615 GGTGCTTCGGTGGGCTCTTGGGCTATGAGATCTTTCAGCAATCATTTGCCCTTTG 2674  
 Db 2659 GGTCTTCGGTGGGCTCTTGGTTCATGAAATTTCTTCAGGAAACACAGCCCTCTG 2718  
 Qy 2675 GATGCGGATGCTGCGCTGTAATTTTGAAGATTTTCTTACATCACTCATCGT 2734  
 Db 2719 GATGCGGATGCTGCGCTGTAATTTTCTGAAAGATTTGTAATCACTCATCGT 2778  
 Qy 2735 GATCCTTGAGACATATCCCTCTGCTTACCTGATACCTGCGCATGTTAT 2794  
 Db 2779 ATACCTTGAGACATATCCCTCTTGGCTATTCACATGCTGCGCATGCTGCT 2838  
 Qy 2795 GACAGGAAATTTATCACTCCAGAGCTGAAATATGTTGCCAGCTGTTGATGCTACT 2854  
 Db 2839 GACAGGAAATTTATCACTCCAGAGCTGAAATATGTTGCCAGCTGTTGATGCTACT 2898  
 Qy 2855 TTTTATCTGCAATTTTGTCTACAGCATCTGAAATGAGATGAGTGTGTTGAATTGA 2914  
 Db 2899 TTTTATCTGCAATTTTGTCTACAGCATCTGAAATGAGATGAGTGTGTTGAATTGA 2958  
 Qy 2915 TGACTGAGTGAAGAAAGACAGTCTGAGTCAATGAGTGTCTCTCAACCTCTTTC 2974  
 Db 2959 TGACTGAGTGAAGAAAGACAGTCTGAGTCAATGAGTGTCTCTCAACCTCTTTC 3018  
 Qy 2975 TGTGTTCAAGGACTCTCAAGGTCATAGCTGTTGATTACAAAGCTTCAACCTGCATC 3034  
 Db 3019 TGTGTTCAAGGACTCTCAAGGTCATAGCTGTTGATTACAAAGCTTCAACCTGCATC 3078  
 Qy 3035 AAAGGTTGAGATGATGAGAGTCTCAAGCTATATACATTCAAATGAGCTACTTAT 3094  
 Db 3079 CAAGGCGGAGAGAGAGAGAGTCTCAAGCTGATACATTCAAATGAGAGCTCTTCT 3138  
 Qy 3095 GATACCTCTTACCACTTGTCTTATTTGAATTCATTTGTTGCTGCTGGGTTTCAA 3154  
 Db 3139 GATACCTCTTACCACTTGTCTTATTTGAATTCATTTGTTGCTGCTGGGTTTCAA 3198  
 Qy 3155 TGGATCAATTAAGGATGAGTCAATGAGGCCCCCTTTGGGAGCTATTTCTTTCAT 3214  
 Db 3199 TGGATCAATTAAGGATGAGTCAATGAGGCCCCCTTTGGGAGCTATTTCTTTCAT 3258  
 Qy 3215 TGGGATGATGCTCATCTTATCCCTTTCAGAGGTTTGTGTAAGGCAAAACAGAC 3274  
 Db 3259 TGGGATGATGCTCATCTTATCCCTTTCAGAGGTTTGTGTAAGGCAAAACAGAC 3318  
 Qy 3275 ACCAAGATTTGATGCTGCTGCTCATTTCTGCTGCTTCTTCTGCTTCTTGGGT 3334

Db 3319 GCCAAGCATTTGTCATTGCTGTGTCATCTCTGCTTCGATCTTCTGCTTGGGT 3378  
 Qy 3335 TCGATGATCTCTTCTCTTGGAGAGATGATGATGCTCTTCTTGAAGAGTGTGTTGCA 3394  
 Db 3379 CCGATTCACCCCTTCTTGGAGAGATGATGATGCTCTTCTTGGAGAGTGTGTTGCA 3438  
 Qy 3395 TTGCACTGATGATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 3427  
 Db 3439 TTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3471  
 RESULT 5  
 AA299491  
 ID AA299491 standard; DNA; 3568 BP.  
 XX  
 AC AA299491;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE DNA encoding a maize cellulose synthase.  
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
 XX transgenic plant; plant breeding marker; ss.  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 63..3236  
 FT /\*tag = a  
 FT /product= "cellulose synthase"  
 FT /note= "no termination codon given"  
 PN WO200009706-A2.  
 PD 24-FEB-2000.  
 XX  
 PR 16-AUG-1999; 99WO-US18760.  
 XX  
 PR 17-AUG-1998; 98US-0096822.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PI Dhurga KS, Helentjaris TG, Bowen BA, Wang X;  
 XX  
 DR WPI; 2000-224343/19.  
 DR P-PsDB; AAY84107.  
 XX  
 PT New genes which encode maize cellulose synthase polypeptides in plants  
 PT useful for modulating the expression of cellulose synthase in plants  
 PT and to produce transgenic plants expressing the novel protein -  
 XX  
 PS Claim 1; Page 82-87; 119pp; English.  
 XX  
 CC The present sequence encodes a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain

CC antibodies specifically immunoreactive with a protein.  
XX Sequence 3568 BP; 879 A; 825 C; 941 G; 922 T; 1 other;  
SQ

Query Match 64.7%; Score 2546.8; DB 21; Length 3568;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

```
QY 177 CACAAACCGAAGAGCTCTGTCATCCCGCCGACGCGGATCCCGGGCCGAAAGCCGCG 236
DB 6 CCGAGCGCTCCGAGCTGTCATCCCGCCGACGCGGATCCCGGGCCGAAAGCCGCAATG 65
QY 237 CCGAGCAGAAACGGGAGGAGTGTGCAAGATTGGGGGACGAGCTGCGCTTCCCGGCG 296
DB 66 GACGACGCGAAGCGGACGAGTGTGCAAGATTGGGGGACGAGCTGCGGCGCAACCCGAC 125
QY 297 GGGGAGCCCTTGTGTGCGTGCACAGTGCCTTCCCGCTGCGCGGAGCTGCTACGAA 356
DB 126 GGGGAGCCCTTGTGTGCGCTGCACAGTGCCTTCCCGCTGCGGAGCTGCTACGAG 185
QY 357 TAGAGCCGCGGAGGAGGACGCAAACTGCCCCAGTGCAGAACTCGATACAGCGCTTC 416
DB 186 TAGAGCCGCGGAGGAGGACGCAAACTGCCCCAGTGCAGAACTCGCTTCAAGCGCTTC 245
QY 417 AAGGCTGCGCAAGTGTGACGCGGTGACGAGAGGAGGAGGAGCGGCTGATGACCTGGA 476
DB 246 AAGGAGTGTGCGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 305
QY 477 GAGTTCAACTGG---GACGCGCATGACTGCGAGTGTGCGCGGAGTGCATGCTTACGCG 533
DB 306 GAGTTCAACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 365
QY 534 CACATGAGTACGCGCGGTGAGAGTACCTTAATGCGCGCGCAAGCTTTCCAGCTCAAC 593
DB 366 CACATGAGTACGCGCGGTGAGAGTACCTTAATGCGCGCGCGCAAGCTTTCCAGCTCAAC 425
QY 594 CCGAATGTTCCACTCTCAACCAACGCGGCAATGAGTGAATGATCCACCGGAGGAGCA 653
DB 426 CCGAATGTTCCACTCTCTCAACCAACGAGTGTGATGATGATCCCGCGGAGGAGCAAC 485
QY 654 GCGCTGTGCTCTTCTTTCATGAGTGTGAGGAGGAGGAGTACATCCCTTCTTATGCG 713
DB 486 GCGCTGTGCTCTTCTTTCATGAGTGTGAGGAGGAGGAGGAGTACATCCCTTCTTATGCG 545
QY 714 GATCCGAGTTACTGTGTGCAACCGGTATGACCCATCCAGAGATCTTGTCTGATAT 773
DB 546 GATCCGAGTTACTGTGTGCAACCGAGTGTATGACCCCTTCCAGAGATCTGCGCGCATAT 605
QY 774 GGGTATGTGATGTGTTGCTTGAAGGAGGAGTGAAGATTTGAAGCAGAGCAAGAGAG 833
DB 606 GGGTATGTGATGTGTTGCTTGAAGGAGGAGTGAAGATTTGAAGCAGAGCAAGAGAGG 665
QY 834 ATGACACGAGCGGAGGAGTATGAGTGTGATGATGATGATGATGATGATGATGATGAT 893
DB 666 ATGACACGAGCGGAGGAGTATGAGTGTGATGATGATGATGATGATGATGATGATGAT 725
QY 894 ATGATGAGCAAGCAACCACTGTCCAGAGAAATTTCCACTTCAACGAGCAATTAAT 953
DB 726 ATGATGAGCAAGCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
QY 954 CCAATATAGATATATATATATATATATATATATATATATATATATATATATATATAT 1013
DB 786 CCAATATAGATATATATATATATATATATATATATATATATATATATATATATATAT 845
QY 1014 GTGATGATCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073
DB 846 GTGATGATCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
QY 1074 TGGTTTGCATGCTTGTGATTTCTTGAATCAATCCAAAGTGTCCATATGAGAGAG 1133
DB 906 TGGTTTGCATGCTTGTGATTTCTTGAATCAATCCAAAGTGTCCATATGAGAGAG 965
QY 1134 ACTTACCTAGACCGGCTGTCACTGAGGTTCCAGAGGAGGAGGAGGAGGAGGAGGAGG 1193
```

```
DB 966 AACTATCTTGAACCGGCTGATGATTAAGTTTGAACAGAGAGGAGGAGGAGGAGGAGGAG 1025
QY 1194 CCAATGATATTTCTTGTGAGTACGAGTGTGATCCCTTAAAGAACTCTTGTGATCAACA 1253
DB 1026 CCGTGTGATTTCTTGTGAGTACGAGTGTGATCCCTTAAAGAACTCTTGTGATCAACA 1085
QY 1254 AATATCTGTATATATATATATATATATATATATATATATATATATATATATATATAT 1313
DB 1086 AATATCTGTATATATATATATATATATATATATATATATATATATATATATATATAT 1145
QY 1314 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
DB 1146 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
QY 1374 AAGAAATGGGTTCTTCTTGTGCAACCGTACATATGAACTCGCGCTCAGAGTGTAC 1433
DB 1206 AAGAAATGGGTTCTTCTTGTGCAACCGTACATATGAACTCGCGCTCAGAGTGTAC 1265
QY 1434 TTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
DB 1266 TTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1325
QY 1494 AGAGCAATGAAAGAGAGTATGAGGAAATTCAGGAGGAGATCAATGCTTATGTTGCCAA 1553
DB 1326 AGAGCAATGAAAGAGAGTATGAGGAAATTCAGGAGGAGATCAATGCTTATGTTGCCAA 1385
QY 1554 GCCGAGAAAGTTCTGTAAGAGAGTGAACATGCAAGTGAACCCCTGCGCTGAGAAAC 1613
DB 1386 GCCGAGAAAGTTCTGTAAGAGAGTGAACATGCAAGTGAACCCCTGCGCTGAGAAAC 1445
QY 1614 AATGTTGCTGATCATCTGGAATGATTCAGGTTCTTGTGCGCAAGCGAGGCTTAC 1673
DB 1446 AATGTTGCTGATCATCTGGAATGATTCAGGTTCTTGTGCGCAAGGAGTGTGCGCATAT 1505
QY 1674 TGTGAGGAGAAATGAATGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1733
DB 1506 TGTGAGGAGAAATGAATGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1565
QY 1734 AACCATCAATGAAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
DB 1566 AACCATCAATGAAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625
QY 1794 AATGCTCATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
DB 1626 AATGCTCATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
QY 1854 AAGGAGCAATGTTTATATATATATATATATATATATATATATATATATATATATATAT 1913
DB 1686 AAGGAGCAATGTTTATATATATATATATATATATATATATATATATATATATATATAT 1745
QY 1914 TTCCCTCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1973
DB 1746 TTCCCTCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
QY 1974 TTTTGTATATCAATGAAGGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2033
DB 1806 TTTTGTATATCAATGAAGGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1865
QY 2034 GATGTGTATTTAAGAGGAGGATTTATATGATGATGATGATGATGATGATGATGATGATGAT 2093
DB 1866 GATGTGTATTTAAGAGGAGGATTTATATGATGATGATGATGATGATGATGATGATGATGAT 1925
QY 2094 CCATCAAGAGCTTGCATGCTGCGCAAGGAGTGTGATGATGATGATGATGATGATGATGATGAT 2153
DB 1926 CCATCAAGAGCTTGCATGCTGCGCAAGGAGTGTGATGATGATGATGATGATGATGATGAT 1985
QY 2154 AGGAGCAAAAGAGAGCTACCAAAACCAAAACAGAGA-----AGAAAGATTAATTT 2207
DB 1986 AGGAGCAAAAGAGAGAGCTACCAAAACCAAAACCTTAATTTGAAGAGATTAAGAACTT 2045
QY 2208 TTCAAGAAAGAGAGAGCAATTCCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2267
```

Dp	2046	TTTAAAGAAAAAGAAATTCAAAGCCCTCGCATATGCTCTTGTAATGATGAAGCCGCT	2105
Qy	2268	CCAGAGCTGAGATGAAAAAGCCGGATTTGTAAATCACAAAAAATTAGAAAAAATTT	2327
Dp	2106	CCAGAGACTGAAAAATGAAAAGGCTAAGATTTGTAATCAACGAAGTTGGAAAAAATTT	2165
Qy	2328	GGCCAACTCTTGTTTTTGTACATCCACATTCCTCGAGATGGATGGAACTTGAAAGCT	2387
Dp	2166	GGCCAGCTTCACGTTTTTTTGTGATCCACACTTCTTGGAATGGTGGAACTTGAAAGCT	2225
Qy	2388	GCAAGCTCTGCTCTCTTTTGGAAAAAGCTATACATGTCATTAGTTGTGTTATGAAGAC	2447
Dp	2226	GCCAGTCAGAGTCTCTCTTCGAGAAGACTATACATGTCATCAGTTGTGATAGAAGAC	2285
Qy	2448	AAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGGATCATGTTACAGAAATATTCTA	2507
Dp	2286	AAAACAGAGCTGGGAAAAAGATATTGGTTGATGTTATGGATCAGTCACAGAAATATTCTT	2345
Qy	2508	ACTGGTTTCAGATGACATGTCATGTTGGGGGTCAATTTACTGATACCTAAAGGTT	2567
Dp	2346	ACTGGGTTTAAAGATGCACTGCCATGGTTGGGGTCAATTTACTGATACCTAAAGGCTC	2405
Qy	2568	GCATTCAAAGGTTCTGCACTCTGCAATCTTTCAGATCGCTTTCACAGAGTCTTCGGTGG	2627
Dp	2406	GCTTCAAAGGTTCCGACCTCTCAATTTTCCGATCGTTTTCACAGAGTCTTCGGTGG	2465
Qy	2628	GCTCTTGGGTCATATTGAGATCTTCTTCGCAATCATTCGCCCTCTTGGTATGGGTAAGT	2687
Dp	2466	GCTCTTGGTTCATATTGAAATTTTGTGTGCAACACATGCCCTCTCTGGTATGGGTAAGT	2525
Qy	2688	GGCGGTCGAAATTTTGGAAAAAGTTTCTCAATCACTCAATCGATCGTATCTTGGACA	2747
Dp	2526	GGTGGACTAAAGTTCCCGGAAAAAGTTTCTGATCAATTACTCAATCGATACCTCTTGACA	2585
Qy	2748	TCTATTTCCCTCTTGGGCTTACCTGATCATATGCTGCACTGCTGTTATTTAGACGGGAAATT	2807
Dp	2586	TCTATCCGCTCTTGGCTATTTGTCATGCAATGCTGCACTGCTGCTGCTGACAGGAAATTT	2645
Qy	2808	ATCACTCCAGAGCTGAATTAATGTTGCCAGCCGTGGTTCATGCACTTTTATCTGCATTT	2867
Dp	2646	ATCAGCGCAGAGCTTTAACAAATGTTGCCAGCCTCGGTTCACTGACCTTTTCACTGCATTT	2705
Qy	2868	TTTGCTACAGACATCTAGAAATGAGATGAGTGGTGGAAATTGATGACTGTGGAGG	2927
Dp	2706	TTTGCTACAGACATCTCGAAATAGAGATGAGTGGTGGACATCGATGACTGTGGAGGA	2765
Qy	2928	AATAGAGAGTCTGGGTCATTTGAGAGTGTGTCCTACACCTCTTGTGCTGCTCCAGGGA	2987
Dp	2766	AACAGAGAGTTTGGGTCATTTGAGAGGCGCTGTCTTCACTCTTGTGCTGCTCCAGGGA	2825
Qy	2988	CTTCTCAAAGTCATAGCTGTGTGTGATCTCAAGCTTCACTGATCAATCAAGAGGTGAGAT	3047
Dp	2826	CTCTCTCAAAGTCATAGCTGTGTGTGATGACACGAGCTTCACTGTGACATCCAAAGGCGAGAC	2885
Qy	3048	GATGAGAGATTCTCAGAGCTATATATACATTCAATATGACCTACCTTATGATACCTCTTAC	3107
Dp	2886	GACGAGAGATTCTCAGAGCTGTATACATTTCAAAATGAGACCTCTCTGATATCTCCGACA	2945
Qy	3108	ACCTTGCTCTTATTTGAACCTTATTTGTTGTGTGCTGTGCTTCAAATGCGATCAATAC	3167
Dp	2946	ACCTTGCTCTCTACGAACTTTCATTTGAGATGGTATGCTGCACTCCAAAGCATCAACAC	3005
Qy	3168	GGATATGAGTATGAGGGGCCCTCTCTTGGGAAAGCATCTTGTGACTTTTGGGTGATTTGTC	3227
Dp	3006	GGATATGAGTATGAGGGGCCCTCTCTTGGGAAAGCTCTTGTGCAATTTTGGGTGATTTGTC	3065
Qy	3228	CATCTTATATCCCTTCTTCAAAAGTTTGGTTGAAAGGCAAAACAGACACCAAGATTGTC	3287
Dp	3066	CATCTTATATCCCTTCTCTCAAGGTCGTTTGGAGGCAAAACAGACCAAGATTGTC	3125
Qy	3288	ATCGTCTGTCATATCTGCTGGCTTCAATCTTCTGCTCTTTGGGTTGGATTTGATCT	3347
Dp	3126	ATTTCTGTCGTCATATCTGCTGGCTTCAATCTTCTGCTCTTTGGGTTGGATTTGATCCCG	3185

QY	3348	TTCTTTGCGAAGATGATGTGCGCTTCTTGAGAGAGTGTGATTGCACTAGAT	3407
Db	3186	TTCTTTGCGAAGATGATGTGCTTCTTGAGAGAGTGTGATTGCACTAGAG	3245
QY	3408	GTCAGTGCATCAGCTCCCC	3427
Db	3246	GTCAGCACGTGAGACTTCCC	3265
RESULT	6		
AAZ58265			
ID	AAZ58265	standard; cDNA; 3786 BP.	
AC	AAZ58265;		
XX			
DT	08-MAY-2000	(first entry)	
XX			
DE	Corn cellulose synthase cDNA clone p0097.cqrad17rc.		
XX			
KM	Corn; maize; cellulose synthase; transgenic plant; ss.		
XX			
OS	Zea mays.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..3499	
FT		/*tag= a	
FT		/partial	
XX			
PN	MO200004166-A2.		
XX			
PD	27-JAN-2000.		
XX			
PR	13-JUL-1999;	99MO-US15871.	
XX			
PR	14-JUL-1998;	98US-0092844.	
XX			
PA	(DUPO ) DU FONT DE NEMOURS & CO E. I.		
XX			
PI	Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;		
PI	Ratalaeki JA, Thorpe CJ;		
DR	WP1; 2000-182431/16.		
XX	P-PSDB; AAY58834.		
PT	Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as		
FT	probes for isolating cDNAs and genes encoding homologous proteins, for		
PT	producing transgenic plants -		
XX			
PS	Claim 9; Page 44-45; 93pp; English.		
XX			
CC	The present sequence is that of cDNA clone p0097.cqrad17rc encoding		
CC	a portion (see AAY58833) of corn cellulose synthase (CS). The cDNA		
CC	clone was isolated from a european corn borer 4 times-infected corn		
CC	stage V9 whorl section cDNA library on the basis of homology to		
CC	Arabidopsis and cotton CS sequences. The invention relates to		
CC	isolated nucleic acid fragments encoding plant CS and to CS		
CC	polypeptides. It also relates to the construction of a chimeric		
CC	gene encoding all or a portion of the CS, in sense or antisense		
CC	orientation, where expression of the gene results in altered levels		
CC	of the CS in transformed host cells. The host cells can be used to		
CC	screen compounds for their ability to inhibit CS activity. CS		
CC	nucleic acids are also useful for producing transgenic plants having		
CC	altered levels of CS, and hence altered levels of fibre. CS may		
XX	also serve as a target for the development of novel herbicides.		
XX			
SQ	Sequence 3786 BP; 846 A; 967 C; 1111 G; 862 T; 0 other;		
Query	March	56.9%; Score 2241; DB 21; Length 3786;	
Match	Local Similarity	81.2%; Pred. No. 0;	
Matches	2691; Conservative	0; Mismatches 580; Indels 42; Gaps 6;	
QY	123	GGAAGTGGAGGAGGAGGATGAGGCGAGCGCGGCTGTGTGCGGCTCCACAC	182



Db	194	GGGCGGGGGGGGGGCGCTGAGATGGAGGGCTAACGGCGGGGCTGGTGGCGGGCTCGGATTAAC	253
OY	183	CGCAACGAGCTCGTGTGATCCGCGCGGACGCGCA-TCCGGGGCGGAAGCCGCGCGGGA	241
Db	254	CGGAACGAGCTGTGTGTATCCGCGCGCACCGCGAGTCGGAGACCGCGGGCGCGCGCG	313
OY	242	GCAGAACGGGCG---AGGTGTGCAGATTGGCGGCGACGAGCTCGGCTTGGCCCGGCG	296
Db	314	GGCGCGCGGGCGGAGGCGCGCGGACAGATTATCGGCGCGACGAGAGTGGGGGTGGCTTGCAC	373
OY	297	GGGGAGCCCTTCGTGTGGCGTGCAGACGAGTGGCGCTTCCCGCTGCGCGGGAATGCTTACGAA	356
Db	374	GGGGAGCCCTTCTGTGGCGTGCAGACGAGTGGCGCTTCCCGCTGCGCGCTGCTTACGAG	433
OY	357	TACGAGCGCCGGGAGGGGACGCGAGAACTGCCCCGAGTCAAGACTGATCAAGAGCGCTC	416
Db	434	TACGAGCGCCGGGAGGGGCTCCGAAACGTGCCGCGAGTGCAGAACCCGCTTACAGAGCTTC	493
OY	417	AAAGGCTGTCCAACTGTGTGACCGGTGACGAGAGAGAGACCGGCTGTGATGACTGTGCAAC	476
Db	494	AAAGGCTGTCCCGGGGTGGCGGGCGACGAGGAGGAGGACCGGCTGTGACGACTGTGAGAGGCG	553
OY	477	GAGTTCAACTGG-----GAGCGCCATGACTGCGAGTCTGTGGCGGAGTCC	521
Db	554	GAGTTCGGGCTGTGCAGAGACGGCGCGCCACGAGGAGCGACCGGAGTATGTCGCGGAGTCC	613
OY	522	ATGCTCTACGGCGCATGTAGCTACGAGCGGTGAGGTGACCCCTAATGGCGCGCACAACT	581
Db	614	ATGCTCAGAGGCGCGAATAGTACGAGCGCGCGGGCGGA-----CGCGCACCCCGGCG	664
OY	582	TTCCAGCTCAACCCCAATGTTCCACTCTCAACCAACCGGCAAAATGATGATGACATCCCA	641
Db	665	TTCAACCCCGGTCCCAACGTGGCGGCTCTCAACCAACCGGCGAGATGGTATGATGACATCCG	724
OY	642	CCGAGGAGACGCGCGTGGTGGCTTCTTCAAG-----GGTGTGGGGGAAAGAGG	692
Db	725	CCGAGGAGACGCGCGTGGTGGCTTCTTCAAGTGAACGCGCGCGGGGGGAAAGAGG	784
OY	693	ATATATCCCTCTCTTATGCGGATCCGAGCTTACCTGTGCAACCCAGGCTTATGGAACCCA	752
Db	785	ATCCACCCGCTCTCTTGGCAGATCCCACTTCCAGTGTGCAACCGAGATCCATGGAACCG	844
OY	753	TCCAAGATCTTGTCTGCATATGTGGTATGTTGTGTTGGAAGGAACGATGGAGAT	812
Db	845	TCCAAGATCTTGGCGGCTTACGGAATATGGACGCTGGCTGGAAGGAGATGGAAGGCG	904
OY	813	TGGAAGCAGACAGAGAGATGACCAAGACGGGGAATGATGGTGGTGTGAT---GAT	869
Db	905	TGGAAGCAGGAAGCAGAGACGCGCTCGAGCATGTCAAGACGAGGGGTGGCGGTATTTGGAT	964
OY	870	GGTGAGATGCTGATCTACCACTAATGATGAAAGCAACAACCTGTCCAGAAAATTT	929
Db	965	GCGCAGATGCAAGTCTCCACTAATGATGAAGTAGGACGACATGTGCCAGAAAATTC	1024
OY	930	CCACTTCATCAAGCCAGATTATTCATATAGATGATTAATCATTAATCCGCTGTGGTT	989
Db	1025	CCTATATCATCAAGCCGAATTATTCCTACAGAGATTAATCGTTATCCGCTGTGGTT	1084
OY	990	TTGGGGTCTTCTTCCACTACCGAGTATGCAATCCGTTGAATGATGACTTTGCTTTGGG	1049
Db	1085	TTGGGGTCTTCTTCCACTACCGAGTATGCAATCCGCGAAGATGCAATTTGGCAATTTGGG	1144
OY	1050	CTCATATCTGTAATGTGAAATCTGTTGCGATCTTGAATTTCTTGACATTAATCCCA	1109
Db	1145	CTCATATCTGTAATGTGAAATCTGTTTGGAGTCTTGAATTTTGATCAAGTTCCCA	1204
OY	1110	AAAGTGTTCCTAATGAGAGAGACTTACCTAGACCGGCTGTCACTAGGTTTGACAAAG	1169
Db	1205	AAAGTGTTCCTAATGAGAGAGACTTACCTAGACCGTGTCACTAAGTTTGACAAAG	1264
OY	1170	GAAAGCCAGCAATCTCAACTTGTCTCAATGAATTTCTTGTACGATCGGTATCCCTTA	1229

Db	1265	GAAGGTCAACCTCTCAGCTTGCTCCAAATCGACTCTTGTGTCAGATCGGTGATCCACA	1322
Qy	1230	AAGGAACCTCTTTGGTGCACAACAATATCTGTTCTATCTATCTCTTCCGTGATATTCCT	1289
Db	1335	AAGGAACCTCCTTGGTGCACAGCGAACACTGTCCTTTCACATCCTTTCTGTGATATATCCG	1388
Qy	1290	GTTGATTAAGTTTCTTGTCTATGTTTCTGATGATGAGTGTGCATAGCTAACGTTTGAACA	1344
Db	1385	GTTGGAAGGTCCTCCTCATGTTTCTGATGATGAGTGTGCATATCTTAACGTTTGAAGA	1444
Qy	1350	TTATCTGAAACATCTGAATTTTGGCAAGAAATGGGTTCTCTTTCTGCAACGGTACATATT	1407
Db	1445	TTGCTGAAACATCTGAATTTTGGCAAGAAATGGGTTCTCTTTCAAGAAAAAGTTTAAATTC	1500
Qy	1410	GAACCTCGCGCTCCAGATGAGTACTTCCACAGAGAATAGACTACTTGTAAAGAACAAGTG	1469
Db	1505	GAGCCTCGTGTCTCGATGATGATCTTCCACAGAGATAGACTACTTGTAAAGAACAAGTT	1566
Qy	1470	GCAGCAAACTTTGTTAGGGAAGAGAGCAATGAAGAGAGTATGAAGAAATTCAGAGTG	1522
Db	1565	GCTGCTCATTTGTTAGGGAAGAGAGCGATGAAGAGAGATACGAGGAATTCAGAGTA	1622
Qy	1530	AGATTCATATGCTTAATGTTGCCAAAGCCCAAGAAAGTTCTCTGAAGAGATGCAATGCA	1589
Db	1635	AGATTCATATGCTTGGTTGCAAAAGCCCAAAAGGTTCTCTGAAGAGATGCAATGCA	1684
Qy	1590	GATGGAACCCCTCGGCTGGAACAATGTTGTTGATCATCTCGAATGATTCAGGCTCTTC	1644
Db	1685	GATGGAACCCCTCGGCTGGAACAACAGTACGCGATCATCTCGAATGATTCAGGATATTC	1744
Qy	1650	CTTGCGCAAGCCGAGAGCCTTGACTGTGAGGGAATGAACCTGCCAGATGGTTATGTT	1709
Db	1745	CTTGCGCAAGATGGGAGTCTGATGTGGAAGAAATGAGTTGCCCTCGCTGGTTATATGC	1804
Qy	1710	TCTGAGAGGAACACACAGAGCTATATACATCATAGAAGAGCTGGTGTATGATGATTC	1766
Db	1805	TCGAGAGAAAAGAGCCCGATTATPACCATCAAGAAAGCGTGGTGCATGAAATGCACGTG	1866
Qy	1770	GTCGAGAGCTCTGCTGTAATAACAATGCTCCATATTGTTTAAACTTGATTTGATGATC	1822
Db	1865	GTCCGTGCTCTGCTGTCTTATCAATGCTGCATACCTATTGAACTTGAATGTATCAC	1922
Qy	1830	TACATCAACAACAGCAAGGCTATTAAGGAACAATGTGTTTATGATGAGACCTTTACTA	1889
Db	1925	TACATCAACAATAGCAAGGCGATTAAGAGGCTATGTTCTTACATGATGAGATCCTTTGGTG	1988
Qy	1890	GGAAGAAAGGTTTGCTATGTCACAGTTCCTCAAAAGATTTGATGGAGTGAATGCCATGAC	1944
Db	1985	GGAAGAAAGTGTGTATGTACAGTTCCTCAGAGGTTGATGTATGACAAAATATAT	2044
Qy	1950	CGATATGTACACGGAATGTGTCTTTTGTGATATCAACATGAAGGTTGAGTGGATTT	2005
Db	2045	CGATATGTACACGGAAGTGTGTCTTTTGTGATCAATCAACATGAAGGTTTGAACGTAAT	2104
Qy	2010	CAGGTCCAATTTATGTTGTAAGTGTGATGATTTTGAAGAGGACGACTTATATGTTAT	2065
Db	2105	CAGGAGCCCAATTTATGTTGTAAGTGTGATGATTTTGAAGAGGACGACTGATATGTTAT	2166
Qy	2070	GATGCCCCCAAAACAAGAGCCACATCAAGGACTTGCAACTGTGGCCCAAGTGTC	2122
Db	2165	GATGCTCTCTTAAAGGAAGAACCCACATCAAGAACTTGCAACTGTGGCCCAAGTGTC	2222
Qy	2130	TTTTCGCTGCTGCTTTGGCAATAGGAAGCAAAAGAACTACCAAAACCCAAAACAAG	2189
Db	2225	CTCTCTGCTGCTGCAAGGAACAGATATTAAGAAAGACATCAAAACCAAGCGAG	2284
Qy	2190	AAGAAAAATTATTTTTCAGAAAAAGAGAACCAATCCCTGCATATGCTCTTGGT	2249
Db	2285	AAGAAAGAAATATTTTTCAGAAAAAGCAAAAACCAATCTCTGCATATGCTTGGT	2344
Qy	2250	GAAATTTGACGAAGCTGCTCCAGAGCTGAGATGAATAAAGCCGGTATTTGAATCAACA	2309
Db	2345	GAAATTTGATGAAGTGTCCAGGTGTGATATTCAGAAAGCCGGAATGTAAATCAACAG	2404



QY 2310 AAAATGAAAAAATTGGCCAACTCTGTTTTGTTATCACTCACTTCGAGAAAT 2369  
 DB 2405 AAACAGAGAGAAATTTGGCAGTCTTGTGTTTTGTGCACTCAACACTTCTTGAAAC 2464  
 QY 2370 GGTGGAACCTTGAAGAGTGCAGTCTGCTCTCTTTGAAAAGAGTATACATGTCATT 2429  
 DB 2465 GAGAGGAGCCTGGAAGAGCGCAAGTCCAGCTTCTCTGAAAGAGCTATACAGTTATC 2524  
 QY 2430 AGTTGATGTTATGAAGACAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCA 2489  
 DB 2525 AGCTGGGAGCTTCAAGACAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCA 2584  
 QY 2490 GTTACAGAAAGATATCTTAACCTGTTTCAAGATGATGATGATGATGATGATGATGAT 2549  
 DB 2585 ATACAGAGAGATATCTTAACCTGTTTCAAGATGATGATGATGATGATGATGATGATGAT 2644  
 QY 2550 TGCATCCGAGAGCGGCTGATTAAGAGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAG 2609  
 DB 2645 TGCATCCGAGAGCGGCTGATTAAGAGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAG 2704  
 QY 2610 CACCAAGTCTTGGTGGGCTTGGGCTTATGAGATCTTCTTCAAGCAATCAATGCTCC 2669  
 DB 2705 CACCAAGTCTTGGTGGGCTTGGGCTTATGAGATCTTCTTCAAGCAAGCACTGCCA 2764  
 QY 2670 CTTTGTATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2729  
 DB 2765 CTTTGTATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2824  
 QY 2730 ATGATGATCTTGGAGATCTATTCCTCTTGGCTTATGATGATGATGATGATGATGATGATGAT 2789  
 DB 2825 ATGATGATCTTGGAGATCTATTCCTCTTGGCTTATGATGATGATGATGATGATGATGATGAT 2884  
 QY 2790 TTAATGACAGAGAAATTTATCACTCCAGAGCTGAATATGATGATGATGATGATGATGATGAT 2849  
 DB 2885 CTGCTACAGGAGAGATTTATCACTCCAGAGCTGAATATGATGATGATGATGATGATGATGAT 2944  
 QY 2850 TCACTTTTATCTGCACTTTTGTCTACAGAGCTTCAAGAAATGATGATGATGATGATGATGAT 2909  
 DB 2945 GCACTTTTATCTGCACTTTTGTCTACAGAGCTTCAAGAAATGATGATGATGATGATGATGAT 3004  
 QY 2910 ATTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2969  
 DB 3005 ATGACACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3064  
 QY 2970 TTTGCTGTCTCAGAGAGCTTCAAGAGCTATGATGATGATGATGATGATGATGATGATGAT 3029  
 DB 3065 TTTGCTGTCTCAGAGAGCTTCAAGAGCTATGATGATGATGATGATGATGATGATGATGAT 3124  
 QY 3030 ACATCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3089  
 DB 3125 ACATCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3184  
 QY 3090 TTAATGATCTCTTACCACTTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 3149  
 DB 3185 CTCTGATTAACCTTCAAGAGCTTCTCTGAACTTCAATGATGATGATGATGATGATGATGAT 3244  
 QY 3150 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3209  
 DB 3245 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3304  
 QY 3210 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3269  
 DB 3305 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3364  
 QY 3270 AGGACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3329  
 DB 3365 AGGACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3424  
 QY 3330 TGGGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3389  
 DB 3425 TGGGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3484

QY 3390 TTGATTCGCACT 3402  
 DB 3485 CTGACGCACT 3497  
 RESULT 7  
 AA299509  
 ID AA299509 standard; DNA; 3813 BP.  
 AC AA299509;  
 XX  
 DT 03-UTL-2000 (first entry)  
 XX  
 DE DNA encoding a maize cellulose synthase.  
 XX  
 KM Maize; cellulose synthase; stalk quality; strand; silage; cellulose;  
 KM transgenic plant; plant breeding marker; ss.  
 XX  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 215..3493  
 FT /\*tag= a  
 FT /product= "cellulose synthase"  
 FT /note= "no termination codon given"  
 XX  
 PN WO200009706-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 XX 16-AUG-1999; 99WO-US18760.  
 PF 17-AUG-1998; 98US-0096822.  
 PR (PION-) PIONEER HI-BRED INT INC.  
 PA Dhuga KS, Helenjaris TG, Bowen BA, Wang X;  
 PI MPI; 2000-224343/19.  
 DR P-PSDB; AAY84113.  
 DR  
 XX  
 PT New genes which encode maize cellulose synthase polypeptides in plants  
 PT useful for modulating the expression of cellulose synthase in plants  
 PT and to produce transgenic plants expressing the novel protein -  
 PS  
 PS Claim 1: Page 129-134; 119pp; English.  
 XX  
 CC The present sequence encodes a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.  
 XX  
 SQ Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 other;  
 Query Match 56.9%; Score 2241; DB 21; Length 3813;  
 Best Local Similarity 81.2%; Pred. No. 0;

	Matches	2691	Conservative	0	Mismatches	580	Indels	42	Gaps	6
QY	123	GGAAGTGAAGGAGAGAGACGATGAGAGCGCGCGGCTGTGTGGCCGAGTCCCAAAAC	182							
Db	194	GGGGGGGGGGGGGGGGCTGAGATGAGAGCTTAGCGGGGGCTGTGGCCGGCTCGATTAAC	253							
QY	183	CGCAACGAGCTGTGTGTCATCCGCTCGCAGCGGCCA-TCCCGGGCCGAAGCCGCGCGGGA	241							
Db	254	CGAAACGAGCTGTGTGTGATCCGCGCCACCGCGAGTGGGAGACCGCGGGGGCGGCGCG	313							
QY	242	GCAGAACGGGGC-----AGGTGGCCAGATTGGCGGCGACGAGCTGTGGCCTTGCCCGCGGC	296							
Db	314	GCGCGCCCGGGCGAGAGCGCCCGTCCAGATATGGCGGACGAGAGTGGGGGTGGGCTTGAC	373							
QY	297	GGGAGACCCCTTCTGTGGCGTGGCAACGAGTGGCGCTTCCCGTCTGGCGGGAGCTGTACGAA	356							
Db	374	GGGAGACCCCTTCTGTGGCGTGGCAACGAGTGGCGCTTCCCGTCTGGCGGGAGCTGTACGAA	433							
QY	357	TACGAGCGCCCGGAGAGGCAACGAGAACTGCCCCAGTGCAGAGACTCGATACAGGCGCTC	416							
Db	434	TACGAGCGCCCGGAGAGGCTTCCAGACGTTGCCCGCAGTCAAGAACCCGCTTACAGAGGCGCTC	493							
QY	417	AAAGGCTGTCCAACTGTGTGACCCGCTGTACGAGAGAGAGAGACGCGCTGTGATGACCTGGAACAAC	476							
Db	494	AAAGGCTGTCCCGCGGGGTGGCCGGCGACGAGAGAGAGAGACGCGCTGTGAGAACCTGTGAGGGC	553							
QY	477	GAGTTCAACTGS-----GACGGCCATGACTGCGAGTCTGTGGCCGAGTCC	521							
Db	554	GAGTTCCGAGCTGTGACGAGACGCGCGCGCCACGAGAGACGACCCGCACTACGTGCGCGAGTCC	613							
QY	522	ATGCTCTTACGGCCACATGAGCTACGCGCGGTGAGGTGACCTTAAATGGCGCGCCACAAAGCT	581							
Db	614	ATGCTCAAGGGCGAGATAGAGCTACGGCGCGCGCGCGAG-----CGCGACACCCCGGC	664							
QY	582	TTCCAAGCTCAACCCCAATGTTCTCACTCTCTCAACCAACGGGCAATATGTGATGACATCCCA	641							
Db	665	TTCAAGCCCGGCTCCCAAGTGTGGCTCTCTCAACCAACGGGCAAGATGTGTGATGACATCCCG	724							
QY	642	CCGGAGCAGACAGCGCTGAGTGCCTTCTTCAATG-----GGTGTGTGGGGAAAGAG	692							
Db	725	CCGGAGCAGACAGCGCTGAGTGCCTTCTTCAATGAGCGCGCGCGGGCGGGGGCAAGAG	784							
QY	693	ATATATCCCTCTTCTTATGCGGATGCCAGCTTACCTGTGCAACCCAGGCTTATGGAACCA	752							
Db	785	ATTCACCGCGCTCCCTTGTGCAAGATCCAAACCTTCCAGTGCACACGAGATTCATGGACCG	844							
QY	753	TTCAAAGATCTTGTGTGCATATATGGTATGTGTGTGTTCTTGGAGGAACGAGTGGAGAT	812							
Db	845	TTCAAAGATCTTGTGTGCATATATGGTATGTGTGTGTTCTTGGAGGAACGAGTGGAGAT	904							
QY	813	TGGAAGCAGACACAAAGAGAGATGACCAAGACGGGGAAATGATGTGTGTGTGTGT---GAT	869							
Db	905	TGGAAGCAGACACAAAGAGAGATGACCAAGACGGGGAAATGATGTGTGTGTGTGTGTGTGT	964							
QY	870	GATGACGATGCTGATCTCACTCAATATGATGAGAGACAACAACACTGTCCAGGAAATTT	929							
Db	965	GGCAGCAGATGCAAGATCTGCCACTAATGATGAACTAGACCACTTGTCCAGAAATGTC	1024							
QY	930	CCACTTCCATCAAGCCAGATTATTCATATAGAGTGAATTATCATTTATTCGGCTGTGTGT	989							
Db	1025	CTATATCATCAAGCCAGATTATTCCTTACAGAGATGATTATCTGTATTCGGGTGGTGTGT	1084							
QY	990	TTGGGGTCTTCTTCCACTACCGAGATGATGCATCCGGTGAATGATGCAATTTGCTTGTGG	1049							
Db	1085	TTGGGGTCTTCTTCCACTACCGAGATGATGCATCCGGTGAATGATGCAATTTGCTTGTGG	1144							
QY	1050	CTCATATCTGTATCTGTGAATCTGTGTTGGCATGTCTTGGATCTTGATCAATTTCCCA	1109							
Db	1145	CTCATATCTGTATCTGTGAATCTGTGTTGGCATGTCTTGGATCTTGATCAAGTTCCCA	1204							
QY	1110	AAAGTGTTCCTTATTTGAGAGAGACTTACTTAACCGGCTGTCTACGTAGGTTGACAAG	1168							
Db	1205	AAAGTGTTCCTTCAATGAGAGAGACTTACTTACCTGAGACCGTTGTGTCACTAAGGTTTGAACAAG	1264							

QY	1170	GAAGCCAGCATCTCCAACTTGGCTCCAAATGATTTCTTTGCTAGTAAAGTTGACCTCTTA	1223
Db	1285	GAAGGTCAACCTCTCAGCTGGCTCCAAATGACCTCTTTGGTCAGTAAGTTGAATCCACA	1324
QY	1230	AAGGAACTCCTTGTGTGCACAAACAATACGTGTTCTATCTATCCCTTTCGGTGAATATCT	1289
Db	1325	AAGGAACCTCCCTTGGTTCACAGGCAACAGTCTCTTCCATCTCTTCTGTGTGATTAATCCG	1384
QY	1280	GTGTATAAGTTTCTTGTATGTTCTGTATGATGTGTCTGCATGCTAACGTTTGAAACA	1349
Db	1385	GTTAGAAAGGTCTCCTCTATGTTTCTGAAGATGTGTCTGCATATGCTTAAGCTTTGAAGCA	1444
QY	1350	TTATCTGAAATCATCTGAATTTTGCAAAGAAATGGGTTCTCTTTCGCAAAACGGTACAAATTT	1409
Db	1445	TTGTCTGAACATCTGTAATTTTGCAAAGAAATGGGTTCTCTTTCAGCAAAAAAGTTTAATATC	1504
QY	1410	GAACCTCGCGCTCCAGAAGTGTACTTCCAAACAGAAATAGACTACTTTGAAAGCAAGGTG	1469
Db	1505	GAGCTCGTGCTCTGAGTGGTACTTCCAAACAGAAATAGACTACTTTGAAAGCAAGGT	1564
QY	1470	GCACCAAACTTTGTTAAGGAGAGAGCAATGAAGAGAGATATGAGAAATTCAAAGTG	1529
Db	1565	GCTCCTTCATTGTTTGAAGAGAGAGGCGCATGAAAGAGAAATACGAGAAATTCAAAGTGA	1624
QY	1530	AGATTCATATGCTTAGTTGCGCAAGCCCAAGAAATGTTCTGAAAGAAAGATGAGCAATGCCA	1589
Db	1625	AGATTCATATGCTTAGTTGCGCAAGCCCAAGAAATGTTCTGAGAAAGATGAGCAATGCCA	1684
QY	1590	GATGGAACCCCTGCGCTCGAAACAATGTTGTGATCATCTGGAATGATTCAGGTCTTC	1649
Db	1685	GATGGAAGCCCTGCGCTCGAAACAAGTACGTCATCATCTGGAATGATTCAGGTATTC	1744
QY	1650	CTTGCGCAAAAGCGAGGCGCTTGACTGAGAGGAAATGAATGCCACGATTTGTTAATGTT	1709
Db	1745	CTTGCGCAAAATGCGCGGTCTGATGTGAAGAAATAGTTGCTCCCTGGTATTATGTC	1804
QY	1710	TCTAGAGAAACGACGAGGCTATTAACCATATTAAGAAAGCTGTGCTATGAATGACATTCG	1769
Db	1805	TCGAGAGAAAGAGGCGGCTTATTAACCATCAACAAGAGGCTGTGCATGAATGACCTG	1864
QY	1770	GTCCGAGTCTCTGCTGTACTPAACAATAGCTCCATATTTGTTAAACTTGAAATGTGATCAC	1829
Db	1865	GTCCGTCTCTGTGCTGTACTPAACAATAGCTGCATTAACCTTAATGAACTTGACCTGTGATCAC	1924
QY	1830	TACATCAACAACACAAAGGCTATAAAGAAAGCAATGTGTTTATGATGAGCACTTTACTA	1889
Db	1925	TACATCAACAATACCAAGGCCATTAAGAAGGCTATGTCTTTCATGATGATCCTTTGGTG	1984
QY	1890	GGAAGAAAGTTTGCTATGTACAGTTCCTCCAAAGATTTGATGGAAATGATCCGCATGAC	1949
Db	1985	GGAAGAAAGTGTCTATGTACAGTTCCTCCACAGGTTTGATGTGATTTGCAAAAAATGAT	2044
QY	1950	CGAATATGTACCCGGAATGTGCTTTTGTGATATCAACATGAAGAAAGTTTGGATGATAT	2009
Db	2045	CGAATATGTACAGAAAGTGTCTTTTGTGACATCAACATGAAGAAAGTTTGGAGCGATTT	2104
QY	2010	CAGGTCCAATTAATGTGTGACTGATATGTATTTATGAAGGACAGCATTAATATGTTAT	2069
Db	2105	CAAGGACCCATTATATGTGGTACTGATATGTGTTTTCAGACGGCAGCACTGTATGTTAT	2164
QY	2070	GATGCCCCCAAAACAAGAGCCCATCAAGGACTTGCAACTGTCTGGCCCAAGTGTGC	2129
Db	2165	GATGCTCTCTTAAAGAAAGAACCCACATCAAGAACTTGCAACTGTCTGGCCCAAGTGTGC	2224
QY	2130	TTTGTCTGTGCTGCTTTGGCAATAGGAGCAAAAGAACTACCAACCCAAACAGAG	2189
Db	2225	CTCTCTTGCTGCTGACAGAGAACAGATATAAAGAAAGAACTACCAAAACCAAGACGAG	2284
QY	2190	AAGAAAAAGTTATTTATTTTCAAGAAAGAAAGAACCAATCCCTGCAATATGCTCTTGCT	2249
Db	2285	AAGAGAAAGAAATTTATTTTTCMAAAGACGAAAAACCATCTCTGTGATATGCTTTGGGT	2344

QY 2250 GAAATTGACGAAGCTGCTCCAGAGAGCTGAGAAATGAAAAGCCGGTAATTGTAATGCAACA 2309  
 DB 2345 GAAATTGAGAAAGTGTCTCCAGGTGTGATATCGAAGAGCCGGAATCGTAATATCAACG 2404  
 QY 2310 AAATTGAAAGAAATTTGGCCCAATCTTGTGTTTGTATACATCCACACTCTCGAGAAAT 2369  
 DB 2405 AAATCTGAGAGAAATTTGGGCGAGCTTCTGTTTTGTGCGATCAACACTCTTGAGAAC 2464  
 QY 2370 GGTGGAACCTTGAGAGAGTCAAGTCTGCTTCTCTTTGAAAGACCTATACATGTCATT 2429  
 DB 2465 GGAAGGACCTGGAAGAGCCGAAGTCCAGGCTCTCTTGAGAGGACCTATACATGTTATC 2524  
 QY 2430 AGTTGTGTTATGAAACAGACAGACTGGGGAAAAAGATTGGCTGATCTATGATCA 2489  
 DB 2525 AGCTGGGCTACCAACAGACGACCTGGGAAAAAGATTGGCTGATCTATGATCA 2584  
 QY 2490 GTTACGAGAGATATCTCAACTGGTTCAAGATGATGATGATGGTGGGGCTCAATTTAC 2549  
 DB 2585 ATACAGAGAGATCTTGACTGATTTAGATGCACTGCAATGGCTGGCGTCTATTTAC 2644  
 QY 2550 TGCATACCTAAACGGGTTGCATTCAAGGTTCTGACCTCTGAACTCTTTCAGATGCTTT 2609  
 DB 2645 TGCATCCCGAAGCGGCTGATTCAAAGTTCTGCGCTCTGAACCTTCCGACGCTTT 2704  
 QY 2610 CACCAAGTGTGGTGGGCTCTTGGGCTATGAGATCTTCTTCAAGCAATCATTTGCCCT 2669  
 DB 2705 CACCAAGTCTTGTGGGCTTGGGCTTGGGCTCGTGAATTTCTTCAAGCAACCTCCCA 2764  
 QY 2670 CTTTGGTATGGATGATGATGGGCTGAAATTTTGGAAAGATTTTCTTCAATCAATCTCC 2729  
 DB 2765 CTTTGGTATGAGATGAGCGGCGGCTTAAATTTCTGAAAGTTTCTTATATCAACTCC 2824  
 QY 2730 ATCGTATCTTGAACATCTATTTCCCTTGGCTTACGTAATTTCCCTGCACTCTGT 2789  
 DB 2825 ATCGTTATCTCGAAGCTCATTTCTCTGCTTACTGTAATTTCTTCTGCTGCACTCTG 2884  
 QY 2790 TTATTGACAGGAAATTTATCACTCCAGAGCTGAATATTTGCGAGCTGTGTTGATG 2849  
 DB 2885 CTGCTCAAGGGAAGTTTATCACACCAAGCTTACCAATGTCCTCCAGTATCTGTTTCA 2944  
 QY 2850 TCACCTTTTATCTGCAATTTTGTGTAAGAGCATCTTGAATGAGATGAGATGATGATG 2909  
 DB 2945 GCACTTTTCTATCTGCACTCTGTAAGCGGCACTCTGAAATGAGATGAGATGAGATG 3004  
 QY 2910 ATTGATGACTGTGAGAGAAATGAGCAATTTGAGGCTATTTGAGAGTGTGCTTCAACCTC 2969  
 DB 3005 ATGACGAGCTGTGAGAGAAATGAGCAATTTGAGGCTATTTGAGAGCGTTTGGGCACTCTG 3064  
 QY 2970 TTTGCTGTGTTCAAGGAGCTTCTCAAGGCTATGAGCTGTGATATACAGACTTCAACCTG 3029  
 DB 3065 TTGCGGCTGTTCAGAGGCTCTGTAAGGCTGTGCGCCGCAATCAACAGACTTCAACCTG 3124  
 QY 3030 ACATCAAGGCTGTGAGATGATGAGAGATCTCAAGAGTATATCATTTCAATGAGATCAAC 3089  
 DB 3125 AGCTCGAAGGCGGAGAGAGAGAGATTTCTCGAGCTGTAACAGTTCAAGTGTGAGACCA 3184  
 QY 3090 TTATTGATACCTCTCAACACTTGTCTTATGAACTTCAATGATGATGATGATGATGATG 3149  
 DB 3185 CTGCTGATACCCCGACCAAGCTCTCTGCTGTAACCTTCAACGAGTGTGAGCGGATC 3244  
 QY 3150 TCMAATGCAATCAATTAACGATATGATGATGAGGCGGCTCTTGGGAAAGCTATTTT 3209  
 DB 3245 TCMAACCGATCAACCAAGGCTATGATGATGAGGCGGCTCTTGGGAAAGCTTCTTTC 3304  
 QY 3210 GCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3269  
 DB 3305 GCCTTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3364  
 QY 3270 AGGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3329  
 DB 3365 AGGACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3424  
 QY 3330 TGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3389

DB 3425 TGGTCCGCTGACCCCTTCTGCGCAAGACAAAGCCGCTCTGAGAGATGTGC 3484  
 QY 3390 TTGATTGCAACT 3402  
 DB 3465 CTGACTGCAACT 3497  
 RESULT 8  
 AA299524  
 ID AA299524 standard; DNA; 3813 BP.  
 XX  
 AC AA299524;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE DNA encoding a maize cellulose synthase.  
 XX  
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
 KW transgenic plant; plant breeding marker; ss.  
 XX  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 215..3493  
 FT /\*tag=a  
 FT /product="cellulose synthase"  
 FT /note="no termination codon given"  
 PN MO200009706-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 16-AUG-1999; 99WO-US18760.  
 XX  
 PR 17-AUG-1998; 98US-0096822.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
 XX  
 DR MPI: 2000-224343/19.  
 XX  
 DR P-PSDB; AY84118.  
 XX  
 PT New genes which encode maize cellulose synthase polypeptides in plants  
 PT useful for modulating the expression of cellulose synthase in plants  
 PT and to produce transgenic plants expressing the novel protein -  
 XX  
 PS Claim 1; Page 168-173; 119pp; English.  
 CC  
 CC The present sequence encodes a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.  
 CC  
 CC Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 other;

Query Match 56.9%; Score 2241; DB 21; Length 3813;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

QY 123 GGAATGAGAGGAGAGAAAGCATGAGAGCGCGCGGCTGTGTGCGCGCTCCCAAC 182  
DB 194 GGGCGGGGCGGGGCGCTGAGATGAGGCTAGCGGGGCTGTGTGCGCGCTCCGATAC 253  
QY 183 CGCAAGAGCTGTGTGTATCCGCGCGAGCGGA-TCCGGGCGGAAGCGCGCGGA 241  
DB 254 CGAAGAGAGCTGTGTGTATCCGCGCGAGCTGAGGAGCGCGCGCGCGCGCG 313  
QY 242 GCAGAGCGGC-----AGGTGCCAGATTGTGCGCGAGCGCTGTGTGCGCGCGC 296  
DB 314 GCGGCGCGGCGAGGCGCGCTGCGAGATATGCGCGAGCGAGTGTGTGTGTGTGCAC 373  
QY 297 GGGAGCCCTTGTGTGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 356  
DB 374 GGGAGCCCTTGTGTGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 433  
QY 357 TACGAGCGCGGAGGAGCGAGAACTGCCCGAGTGCAGAACTGCATCAAGCGCTC 416  
DB 434 TACGAGCGCGGAGGAGCGCTGCGAGCGCTGCCGAGTGCAGAACCGCTCAAGCGCTC 493  
QY 417 AAGGAGT 476  
DB 494 AAGGAGT 553  
QY 477 GAGTTCAACTGT-----GACGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 521  
DB 554 GAGTTGT 613  
QY 522 ATGCTTACGCGCAATGAGCTACGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 581  
DB 614 ATGCTTACGCGCAATGAGCTACGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 664  
QY 582 TTCAAGCTCAACCCCAATGTTCATCTCTCAACCAAGGCAATGTGTGTGTGTGTGTGTGT 641  
DB 665 TTCAAGCCCGTCCCAATGT 724  
QY 642 CCGGAGCGAGCGCGGT 692  
DB 725 CCGGAGCGAGCGCGGT 784  
QY 693 ATATATCCCTTCTTATGCGAATCCAGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752  
DB 785 ATCAACCGCTCCCTTGT 844  
QY 753 TCAGAGATCTTGT 812  
DB 845 TCAGAGATCTTGT 904  
QY 813 TGGAGCGAGAGAGAGAGATGTCAAGAGCGGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 869  
DB 905 TGGAGCGAGAGAGAGAGATGTCAAGAGCGGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 964  
QY 870 GGTAGCGATGT 929  
DB 965 GGTAGCGATGT 1024  
QY 930 CCACTTCATCAAGCGAGATTAATCATATAGATGATTAATCATTAATCGCTGTGTGTGT 989  
DB 1025 CCACTTCATCAAGCGAGATTAATCATATAGATGATTAATCATTAATCGCTGTGTGTGTGT 1084  
QY 990 TTGGGAGTCTTCTTCACTACGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1049  
DB 1085 TTGGGAGTCTTCTTCACTACGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1144  
QY 1050 CTCAATCTGTATCTGTGAATCTGT 1109  
DB 1145 CTCAATCTGTATCTGTGAATCTGT 1204

QY 1110 AAGTGTCCCTATTTAGAGAGACTTACTAGACCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1169  
DB 1205 AAGTGTCCCTATTTAGAGAGAGACTTACTAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1264  
QY 1170 GAAAGCGAGCATCTCAATGT 1229  
DB 1265 GAAAGCGAGCATCTCAATGT 1324  
QY 1230 AAGAACCTCTTGT 1289  
DB 1325 AAGAACCTCTTGT 1384  
QY 1290 GTTGTATAGGTTCTGT 1349  
DB 1385 GTTGTATAGGTTCTGT 1444  
QY 1350 TTATCTGAACATCTGAATTTGCAAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1409  
DB 1445 TTGTCTGAACATCTGAATTTGCAAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1504  
QY 1410 GAACCTGCGCTCCAGAGT 1469  
DB 1505 GAGCTGT 1564  
QY 1470 GCAGCAAACTTGT 1529  
DB 1565 GCTGTCTTCAATTTGT 1624  
QY 1530 AGAATCAATGCTTGT 1589  
DB 1625 AGAATCAATGCTTGT 1684  
QY 1590 GATGAAACCCCTGT 1649  
DB 1685 GATGAAACCCCTGT 1744  
QY 1650 CTTGCGCAAAAGCGAGGCTTGT 1709  
DB 1745 CTTGCGCAAAAGCGAGGCTTGT 1804  
QY 1710 TCTAGAGAGAAAGAGAGGCTATTAACCATATTAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1769  
DB 1805 TCTAGAGAGAAAGAGAGGCTATTAACCATATTAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1864  
QY 1770 GTCCAGTCTGT 1829  
DB 1865 GTCCAGTCTGT 1924  
QY 1830 TACATCAACAAGCAAGGCTATTAAGAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1889  
DB 1925 TACATCAACAAGCAAGGCTATTAAGAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1984  
QY 1890 GGAAGAAAGT 1949  
DB 1985 GGAAGAAAGT 2044  
QY 1950 CGATATCTCAACGGAATGT 2009  
DB 2045 CGATATCTCAACGGAATGT 2104  
QY 2010 CAGGCTCAATTTATGT 2069  
DB 2105 CAGGCTCAATTTATGT 2164  
QY 2070 GATGCGCCCAAAAGAGAGAGCAATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2129  
DB 2165 GATGCGCCCAAAAGAGAGAGCAATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2224  
QY 2130 TTTTGT 2189  
DB 2225 CTGTCTGT 2284  
QY 2190 AAGAAAGATTTATTTTTCAGAGAAAGAGAGCAATCCCTGTGTGTGTGTGTGTGTGTGTGTGT 2249

Db	2285	AAGAGAAAAGATTATTTTTCAGAAAGCAGAAAACCATCTCTGCATATGCTTGGT	2344
Qy	2250	GAAATTGACGAAGCTGCTCCAGAGCTGAGATGAAAAGCCGGTATTGTAAATCAACA	2309
Db	2345	GAAATTGATGAAAGGTGCTCCAGAGTGTGATATCGAAGAGCCGGAAATGTAAATCAACAG	2404
Qy	2310	AAATTAGAAAAAGAAATTTGGCCAAATCTTCTGTTTTTGTATATCACTCACTTCTGGAAT	2365
Db	2405	AAACTTAGAAGAAATTTGGGCACTCTTCTTTTTTGTGCGATCAACACTTCTTGAGAAC	2464
Qy	2370	GGTGAACCTTGAAGAGTGCAAAGTCCGTCTCTCTTTGAAAGAGCTATACATGTCACT	2429
Db	2465	GGAGGGACCTCGAAAGAGCGCAAGTTCAGCTTCTCTTCTGAAGAGACTATACATGTTATC	2524
Qy	2430	AGTTGTGTTATGAAAGCAGACAGACTGGGGAAAAAGATTTGGCTGATCTATGATCA	2489
Db	2525	AGTCGGGCTACGAAGACAGACCGACTGGGGAAAAAGATTTGGCTGATTTACGAGATCG	2584
Qy	2490	GTTACAGAAATATTTCTACTGCTTTCAAGATGCAATGTCAATGCTGGCCGTCAATTTAC	2549
Db	2585	ATCACAAGGATATCTTGAATCTGGAATTTAAGATGCACTGCATATGCTGGCGGTCTAATTAAC	2644
Qy	2550	TGCATACCCTAAACGGGTGTCATTCAAAGTTCTGCACCTCGAATCTTTCAGATGCTT	2609
Db	2645	TGCATCCGGAAGCCGCTGCATTCAAAGTTCTGGCCCTCTGAACTTTCCGACGCTTT	2704
Qy	2610	CACCAAGTCTCGGTGGGCTCTTGGGTCTATTTAGATCTTCTTCAGCAATCATTTGCCCT	2665
Db	2705	CACCAAGTCTTGGTGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAACATCTGCCCA	2760
Qy	2670	CTTTGGTATGGGATATGTGGCGGTCTGAATTTTGGAAAGATTTTCTTACATCACTCC	2725
Db	2765	CTTTGGTATGGGATATCGGCGCGGCTAAATTTCCGGAAGATTTTCTTATATCACTCC	2824
Qy	2730	ATCGGTATCCTTGGACATCTATTCCTCCCTGTGGTTACTGACATTTGCCGCATCTGT	2789
Db	2825	ATCGTTATTCCTTGGACATCTATTCCTCCCTGTGGTTACTGACATTTGCCGCATCTGT	2884
Qy	2790	TTATTTGACAGGAAATTTTACCTCTCAAGCTCAAGCTGAATATGTTCAGCGCTGTGTTCAATG	2849
Db	2885	CTGCTCAAGGGAAGTTTATCAACACCAAGCTTACCAATGTCCGACGATCTGTGTTCAATG	2944
Qy	2850	TCACCTTTTATCTGCATTTTGTCTACGAGCATCTTGAATGAGATGAGTGTGTTGGA	2909
Db	2945	GCACCTTTTATCTGCATTTTGTCTACGAGCATCTTGAATGAGATGAGTGTGTTGGA	3004
Qy	2910	ATTATGATCTGTGAGGAATGAGAGATTTCTGGGTCAATTTGAGAGTGTCTCTCACACTC	2965
Db	3005	ATCAACGACTGTGTGAGGAAGAGAGATTTCTGGGTCAATTTGAGAGTGTCTCTCACACTC	3064
Qy	2970	TTTGTGTGTTTCCAGGGACTCTCAAGGTCAAGCTAGCTGTGTTGATACAAAGCTTCAACGTG	3029
Db	3065	TTTGTGTGTTTCCAGGGACTCTCTGTGAGGTGTTCGCCGACATGACACAGCTTCAACGTG	3124
Qy	3030	ACATCAAGGGTGAAGATGAGAGATTTCTCAGAGCTATATACATTCAAATGACCTAAC	3089
Db	3125	ACGTCAAGGGCGGGAGAGAGAGAGATTTCTCAGAGCTATATACATTCAAATGACCTAAC	3184
Qy	3090	TTATTTGATACCTCTTACCACTTGTCTTATATTTGAACTTCAATTTGTGTGTGCTGCGCT	3149
Db	3185	CTGTGTATACCTCTTACCACTTGTCTTATATTTGAACTTCAATTTGTGTGTGCTGCGCT	3244
Qy	3150	TCAATATGATCAATTAACGATATGAGATGAGGGGCCCTCTTTGGAGGCTATCTTT	3209
Db	3245	TCGAATGATCAATTAACGATATGAGATGAGGGGCCCTCTTTGGAGGCTATCTTT	3304
Qy	3210	GCATTTTGGGTATGTCTCATCTTATCCCTTTTCAAGAGTTTGGTTGGAAGCGAAAC	3265
Db	3305	GCCTTTTGGGTATGTCTCATCTTATCCCGTTTCAAGAGTTTGGTTGGAAGCGAAAC	3364
Qy	3270	AGGACCAACGATTTGATGTGTGTGCTGATCTGTGCTTCAATCTTCTCGCTCTT	3329

Db	3365	AGGACGCCGACGATCGTCATGTCGTGACATCTGCTGGAGCTCGATCTTCGCTCCTG	3422
Oy	3330	TGGGTTTGGGATGTATCTCTTTCTCTTGCCAGAGATGATGTCGCTTCTTGAGAGATGTGT	3388
Db	3425	TGGGTCGCGCGCGTCGACCGGTTCTCTGCGCAGAGCAAGCAACGCGCGCTCTTGAGAGAGTGTGCG	3488
Oy	3390	TTGGATTGGCAACT	3402
Db	3485	CTGGACTGCAACT	3497
RESULT 9			
ID	ABZ14725	standard; DNA; 3255 BP.	
XX	ABZ14725;		
AC	ABZ14725;		
DT	21-JAN-2003	(first entry)	
XX			
DB	Arabidopsis thaliana stress regulated gene SEQ ID NO 2530.		
XX			
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.		
OS	Arabidopsis thaliana.		
XX			
XX	WO200216655-A2.		
XX			
PD	28-FEB-2002.		
XX			
PF	24-AUG-2001; 2001WO-US26685.		
XX			
FR	24-AUG-2000; 2000US-227866P.		
PR	26-JAN-2001; 2001US-264647P.		
PR	22-JUN-2001; 2001US-300111P.		
XX			
PA	(SCRI ) SCRIPPS RES INST.		
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Harper JF, Kreps J, Wang X, Zhu T;		
XX			
XX	WPI; 2002-304127/34.		
PT	Identifying a stress condition to which a plant cell has been exposed		
PT	and producing plants with increased tolerance to these abiotic stresses		
PT	-		
PS	Claim 144; SEQ ID NO 2530; 577bp + Sequence Listing; English.		
XX			
XX	The invention relates to identifying a stress condition to which a plant		
CC	cell has been exposed, comprising:		
CC	(a) contacting nucleic acid representative of expressed polynucleotides		
CC	in the plant cell with an array or probes representative of the plant		
CC	cell genome; and		
CC	(b) detecting a profile of expressed polynucleotides in the plant cell		
CC	characteristic of a stress response. The method is useful in the		
CC	production of transgenic plants, cells and seeds and in producing plants		
CC	with increased tolerance to abiotic stresses. The present sequence is that		
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used		
CC	in methods of the invention.		
CC	Note: The sequence data for this patent is not represented in the printed		
CC	specification but is based on sequence information supplied to Derwent by		
CC	the European Patent Office.		
XX			
XX			
XX	Sequence 3255 BP; 841 A; 641 C; 813 G; 960 T; 0 other:		
Query Match	38.3%; Score 1509; DB 24; Length 3255;		
Best Local Similarity	67.6%; Pred. No. 0;		
Matches 2207; Conservative	0; Mismatches 1030; Indels 30; Gaps 5;		
Oy	144	ATGAGGCGAGCGCGCGGCTGTGTGTCGCGCGCTCCACAAACCGCAACGAGCTGTGTCATC	203
Db	1	ATGGAACACCGTGTGCTGGTTAATGACCGCGTTCTCACAACAAGAAATGATTTGTCTCAATT	60

OY	204	GC	CGCGG	CGAGG	AGTCCCGGG	CCGAGG	CGCGCGGG	AGGAA	CGGGC	AGGTG	CCAG	263
Db	61	AATG	CCGATG	AGAAATG	CCCGAAT	AGATCA	GTCCAA	GAGCTG	AGTG	GCAC	ATGTTCAA	120
OY	264	ATTG	CGCGG	CGAGG	AGTCCCGGG	CCGAGG	CGCGGG	AGGAA	CGGGC	AGGTG	CCAG	323
Db	121	ATCTG	CAAGATG	AGATG	CAATG	ATGACTG	TGTA	TGGA	GAAACG	CTTTGT	TGGCATG	180
OY	324	TGCG	CTTTCCCG	GTCTG	CGGGAG	CTGCTA	CGAAT	TGAG	CGCGGG	AGGGCA	CGGAAC	383
Db	181	TGTG	CATTTCCG	TGTGT	GTAGAC	CTTGCT	ATAG	GTACG	AAAGCG	AGGCAAT	CAACGT	240
OY	384	TG	CCCCCG	AGTCCAA	ACTCG	ATAC	AGAGCG	CTCA	AGGGCT	CGCA	CGTGTAC	443
Db	241	TGTC	CAAGTCC	AAAAAC	CCGTTTCA	AAAGCT	CTTAA	GGAAGTCC	AAAGG	ATGTA	AGGTAT	300
OY	444	GAG	AGAGAG	AGAGCGG	GTGATG	ACTTG	AGCA	CGAGTTC	CAACTG	GGAGCGG	CAATG	503
Db	301	GAA	AGAGAG	AGATG	ACATG	ATGATTA	TGAC	ATGAGTTGA	---GTATG	AAATAT	ATAGG	357
OY	504	CAGT	CTGTG	CGCGAGTCC	ATGCTCTA	CGGCA	CATG	AGCTAC	CGGCGG	TGAGGTG	ACCT	563
Db	358	ATTG	GAATTTG	ATCA	GGTTTCTG	AAAGATAT	ATGTC	ATCTCG	ATCG	CAATCCG	GTATCCCA	417
OY	564	AATG	CGCGG	CGCCACA	AGCTTTCA	AGCTCA	ACCCCA	TGTTCCA	CTCTCA	CCAA	CGGGCA	623
Db	418	CAAT	CTGATTTG	ATTCAG	CTCCAC	CTGAGCTCTC	AGATTC	CAATGCT	CACTTAA	CGG	CCAC	477
OY	624	ATG	GTGATG	ATCATCC	CAACG	AGAGAC	AGCGC	---TGGT	GCCTCTT	TA	TGGTGGT	680
Db	478	GAG	ACGTTG	AGATTTCTT	CTGTG	ATGACATG	CTCTTAT	TGTTCTT	CTCTTCA	CTTG	TGTGT	537
OY	681	GGG	GGAAA	GAGATP	ATCATCC	CTTCTTAT	GGATGCC	AGTTAC	CTGTG	CAAC	CCGAG	740
Db	538	CATG	GCATAT	AGATTG	ATCATCTG	TTTCTCTT	CTGAC	CCCGAC	CGTGG	CTGC	ACATCA	597
OY	741	TCT	ATGACCC	ATCCAG	ATCTTG	CTGATATG	GGTATG	GTATG	GTGTG	CTTG	GAA	800
Db	598	CTT	ATGTG	ATCTCA	AAAAAGATCTT	CGGTTTAT	GTATG	GAAGTGTG	CGCTTG	GA	AAAGAT	657
OY	801	CG	ATGAG	GAATTTG	GAACAG	AGCA	GAAGAG	ATGCA	CCAG	CGGGG	-----AT	851
Db	658	CG	ATGAG	GAATTTG	GAACAG	AGCA	GAAGAG	ATGCA	CCAG	CGGGG	-----AT	851
OY	852	GATG	TGTG	TGATG	ATGATG	ATGATG	ATGATG	ATGATG	ATGATG	ATGATG	ATGATG	911
Db	718	GAT	CTCG	ATTTTGA	AGATG	GTGATG	ATGCTG	ATTTTCCA	ATGATG	ATGATG	ATGATG	777
OY	912	CA	CTGTC	CAAGAAA	ATTC	CACTTC	ATCA	AGCCAG	ATTAAT	TCATAT	AGATAT	971
Db	778	CC	ATTTGT	CTA	GAAGATAT	CAATCA	AAATCG	AGCA	GAATAAAT	CTTTA	CCGATG	837
OY	972	A	TATAT	CGGCTTG	TGTTGG	GGTCTTTT	TCCA	CTAC	CGAG	ATGAT	TCGAT	1031
Db	838	G	TGCTAC	GTCTG	ATCTT	GTGCTCTT	CTTCA	CTAC	CGATAT	CTTCA	CCCGTCAA	897
OY	1032	GATG	CA	TTGCTTTG	TGCTCA	TATCTG	TATCTG	TAAATCTG	GTG	TG	GCATG	1091
Db	898	GATG	CA	TATGCTTTG	TGCTCA	TATCTG	TATCTG	TAAATCTG	GTG	TG	GCATG	957
OY	1092	A	TTCCTG	ATCA	ATTC	CCAAAG	TGTTCC	CTATG	AGAG	AGAC	CTTAC	1151
Db	958	G	TTCCTG	ATCA	ATTC	CCAAAG	TGTTCC	CTATG	AGAG	AGAC	CTTAC	1151
OY	1152	TC	ACTG	AGGTTG	AGAGAG	GGCCAG	CGATCT	CAACTT	GCTCCA	ATTG	ATTTCTT	1211
Db	1018	TC	ACTG	AGGTTG	AGAGAG	GGCCAG	CGATCT	CAACTT	GCTCCA	ATTG	ATTTCTT	1211
OY	1212	AGT	ACG	GTGAT	TCCCTT	TAAAGAA	CTCCTT	TGTC	CA	CAAAAT	ATCTG	1271
Db	1078	AGT	ACG	GTGAT	TCCCTT	TAAAGAA	CTCCTT	TGTC	CA	CAAAAT	ATCTG	1271
OY	1272	CTT	TCG	GTGAT	TCCCTT	TAAAGAA	CTCCTT	TGTC	CA	CAAAAT	ATCTG	1331

Db	1138	CTTGCTGTGATTTATCCTGTGATPAAGGTCCTTGTTACGATCTGATGATGGTGCCT	1197
Qy	1332	ATGCTAAACGTTGAAGCATATATCTGAACATCTGAATTTGCAAGAAATGGGTCCTTC	1391
Db	1198	ATGCTTACTTTTGAAGCTCTTCTGAGACCGCTGAATTCGAGAGAAATGGGTCCTTC	1255
Qy	1332	TGCAACCGGTACAAATTTGAACCTCGCGCTCCAGATGGTATCTTCCACAGAGATGAC	1451
Db	1258	TGCAAGAAATTTGATATGAGCCTCGCTCCCAATGGTATTTCTCGCATPAAATGAC	1317
Qy	1452	TACTTGAAGAACAAGGTGGCAGCAACCTTGTGGGAGAGACAGCAATGAAGAGAG	1511
Db	1318	TACTTGAAGATPAAAGTTTCATCCGCAATTTGTAGGAGCGGCAAGCATGAAGAAT	1377
Qy	1512	TATGAGAAATCAAGGTGAGAAATCAATGCTTAGTTGCAAGCCCAAGAAAGTTCCTGAA	1577
Db	1378	TATGAGAAATCAAGGTAAAGATCAATGCTTTAGTACAAACAGACAGAAAGTCTGAG	1433
Qy	1572	GAAAGATGACAATGACAATGAAACCCCTGCGCTGGAACAAATGTTGATCATCTCT	1631
Db	1438	GATGGTTGGAATATGACAAGCGGTACACTTGGCCCGGTAATAGTGTGAGATCATCTCT	1497
Qy	1632	GGATGATTCAGGTCCTTCTGCGCAAGCGGACCTTGACTGTGAGGGAATGAACTG	1697
Db	1498	GGCATGATTCAGGTCCTTCTGGAAGTGAACGGTGTTCGTATGTGCAAAACAGAGTTG	1557
Qy	1692	CCACGATTTGTTTATGTTTCTAGAGAGAAAGACAGGCTATTAACATCATAGAAAGCT	1751
Db	1558	CCTGATTAAGTTTACGTTTCTCTGAGAGAGAACCCGATTTGATCACATTAAGAGCT	1617
Qy	1752	GGTGCTATGAATGCAATGGTCCGAGCTCTGCTGTACTTAACAAATGCTCCATATTTGTA	1811
Db	1618	GGAGCTATGAATTCCTCTGATACAGAGCTCTGGGGTTCATCAAAATGCTCTTACTCTG	1677
Qy	1812	AACCTTGATTTGTGATCACTATATCAACAACGCAAGCTATPAAAGAAAGCAATGTTTT	1871
Db	1678	AATGTCGATTTGTGATCACTATATCAACAAATGCAAAAGCTCTTAAGAAAGCAATGTTTT	1737
Qy	1872	ATGATGCAACCTTTACTAGAAAGAAAGGTTTGCATGACAGTTCCCTCAAGATTTGAT	1931
Db	1738	ATGATGATCCTTCACTAGTCAAGAAAGAAATCTGTATGTTCAGTTCCCTCAAGGTTGAT	1797
Qy	1932	GGGATTTGATGCGCATGACGATATGCTAACCGAATGTGTCTTTTGTGATATCAACATG	1991
Db	1798	GGGATTTGATGCGCATGACGATATGCTAAATGCAATGTGTGTCTTTGATATCAATATG	1855
Qy	1992	AAAGTTTGTGATGCTATTCAAGGTCCAATTTATGTTGCTACTGATGTGTATTTAGAGG	2051
Db	1858	AAAGTTTGTGATGCGCTATCAAGGCGCTATATACCTGGTATCAAGGTGTGTTCAGAGG	1917
Qy	2052	CAGGCAATTAATGATTTATGATGCCCCCAAAACAAGAACCAACATCAAGACCTTGCAC	2111
Db	1918	CAGGCGCTTAACGATTTTATGATGACACCGAAGAAAGAGGGCCACGTAACATGTCAT	1977
Qy	2112	TGCTGCGCCAAAGTGTGCTTTGCTGTGTGCTTTGTGCAATAGAAAGCAAAAGAGACT	2177
Db	1978	TGCTGCGCCAAATGTGTGTCTCTATGTTTTGTGTTCAAGAAAGATCTTAAGCAAGACA	2037
Qy	2172	ACCAAAACCAAAACAGAGAGAGAAAGTATTTATTTTCAAGAAAGAAAGAACCAATCC	2231
Db	2038	GTGCGTGGGATTAAGAAAGAAAGAA-----TAGGAAAGCTCAAGACAGATC	2085
Qy	2232	CTGCGATATGCTCTTGTGTGAATTGACGAAGCTCTCCAGAGACTGAGAAATGAAGAGCC	2291
Db	2086	CACCGATTTAGAAATATCGAAGAGGCGCGCTCATAAAGGTTCTAACGTAGAACAGTCA	2145
Qy	2292	GGTATTTGAATCAACAAATATTTAGAAAGAAATTTGGCCCAATCTTCTGTTTTGTACA	2351
Db	2146	ACCGAGGCAATGCAAAATGAAATTTGGAGAGAAATTTGGGCAAGTCTCTCTATTTGTGCA	2205
Qy	2352	TTCACACTTTCGAGATGTGTGAACCTTGAAGAGTCAAGTCTGCTCTCTTTTGAAG	2411





Db 129 TCGAATTACCTGTTGATGAGAACCCGTTGTGGCACTGTAAGCATGTGCAATTCCTGTGT 188  
Qy 340 GCCGGACCTGTAACGAATATACGACGCCGGAGGGGACGCAAGACTGCCCACTGACAGA 399  
Db 189 GTAGACCTTGCTATGATGATACGAAAGACGAATCAAGCTTGTCCACAGTGCAGAAA 248  
Qy 400 CTGATACAGAGGCTCAAGGGCTGCCAAGCTGTGACCGGTGACAGAGAGAGAGCGCG 459  
Db 249 CCGGTTCAACGCTCTTAAGAGAGTCCAGAGTTGAAGTGATGAAAGAGAGATGCA 308  
Qy 460 TCGATGACCTGGAACAAGATTCAACTGGGACGGCATGACTCGCAGTCTGTGGCCAGT 519  
Db 309 TTGATGATTTAGACATAGATTGGA--GTATGAAATTAATGGGATTGGATTGATCAGG 365  
Qy 520 CCATGCTTACGGCCACATGAGCTACGGCCGTGGAGTGAACCTTAATGGCCGCCACAG 579  
Db 366 TTTTGAGAGTATGTCATCTCTGTCGCACTCGGTTTCCCAAACTGATTTTGATTC 425  
Qy 580 CTTTCCAGCTCAACCCCAATGTTCCACTCTCTCAACCAAGGCAAAATGGTGAATGACATCC 639  
Db 426 CAGCTCCACCTGGCTCTCAGATTCATTTGCTGACTTACGGCGACGAGGACGTTGAGATT 485  
Qy 640 CACCGAGACACACCGCG--TGTGCTCTTTCAATGGGTGTGGGGAAAGAGATAC 696  
Db 486 CTTCTGATGACATGCTCTTAATGTTCTCTCTTCACTTGTGTGTCATGGCAATGAGTTC 545  
Qy 697 ATCCCTCTCTTAATGCGGATCCAGCTTACCTGTGCAACCCAGGCTATGACCCATCCA 756  
Db 546 ATCTGTTTCTCTTCTTCAACCCGACCGTGGCTGCACATCGAAGCTGATGTAATCTCAGA 605  
Qy 757 AGGATCTTGTGCAATATGGGTAATGGTATGTTGCTTGGAGAGAACGATGAGAAATTGGA 816  
Db 606 AAGATCTTGGGTTATATGTTATGGAAGTGTGCTTGGAAAGATCGATGAGAGAAATGGA 665  
Qy 817 AGCAGAGCAAGAGAGATGACACGACGGGGA-----ATCATGTGTGTGTGT 867  
Db 666 AGAGAAAGCAAGATGAGAAACTTCAGTGTGTTAGGCATGAAGAAATCTGAATTTGAA 725  
Qy 868 ATGGTGAAGATGCTGATCTACCACTAATGATGAGAAAGAACAACTGTCCAGAGAAA 927  
Db 726 ATGTGATGATGCTGATTTTCCAAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 785  
Qy 928 TTCCACTTCAATCAAGCCAGATTAATCCATATAGGATGATTATCAATTTGCGCTTGTGG 987  
Db 786 TACCAATCAAAATCGAGCAAGATTAATCTTACCAGATGTTAATGTGTACGTCTTGTGA 845  
Qy 988 TTTTGGGTTCTTCTTCACTACGAGTGAATCCGGTGAATGATGATGATTTGCTTTGT 1047  
Db 846 TTTCTGTCTCTTCTTCACTACCGTATTTCTTCAACCCGTCAMAGATCATATGCTTTGT 905  
Qy 1048 GGCTCATATCTGTTATCTGTGAAATCTGTTTGGCATGCTTGAATCTTGAATCAATTC 1107  
Db 906 GGGTATTTCTGTTATATGAGATATGAGATATGGTTGCTTTTCAATGGGTTCTTATACGTCC 965  
Qy 1108 CAAATGAGTTCCCTTATGAGAGAGACTTACTGACCGGCTGTCACTGAGGTTTGCACA 1167  
Db 966 CTAATATGTAACCTATGAG 1025  
Qy 1168 AGAAGGCCAGCATCTCAACTGTGCTCAATTTGATTTCTTTGTCAATGCGTTGATCCCT 1227  
Db 1026 AAGAAAGGAAACCGTCGGGACTATCCCTGTGAGATGTTGTTGATGATGATGATCAT 1085  
Qy 1228 TAAAGAACTCTCTTGTGCTACAAATACTGTCTTATCTTATCTTCCGTGATATTC 1287  
Db 1086 TGAAGAACCTCTCGCTATATCTGCAAAATAGTCTTCTATATCTCTGTGATATTC 1145  
Qy 1288 CTGTGATAGAGTTTCTGTATGTTCTGATGATGATGATGATGATGATGATGATGATGATG 1347  
Db 1146 CTGTGATAGAGTTTCTGTATGTTCTGATGATGATGATGATGATGATGATGATGATGATG 1205  
Qy 1348 CATTAATCGAAACATCTGAATTTGGCAAGAAATGGGTTCTTTCTGCAACCGTACAAATA 1407

Db 1206 CTCTTTGAGAACCGCTGATATCCAGAGAAATGGGTTCTTTCTGCAAGAAATTTCTA 1265  
Qy 1408 TTGAACCTCGCGCTCCAGAGTGTACTTCCAAAGAAAGATAGACTACTGTAAGAGACAGG 1467  
Db 1466 TTGAGCTCGTGTCCGAAATGATATTTCTGCACTAAATATGACTACTTGAAGAAATGAAG 1325  
Qy 1468 TGGAGCAAACTTTGTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527  
Db 1326 TTCAATCCGCAATTTGTTAGGAGAGCGGAGCCATGAAAGAGAGATTAAGAAATTTCAAG 1385  
Qy 1528 TGAGAAATCAATGCTTATGTTGCAAAAGCCAGAAAGTTCTCTGAAGAGATGACAAATGC 1587  
Db 1386 TAAAGATCAATGCTTATGATGCAACACACAGAAAGTGCCTGAGAGATGTTGACTATGC 1445  
Qy 1588 AAGATGAACCCCTGCGCTGGAAGCAATGTTGTGATCATCTGGAATGATTCAGGCT 1647  
Db 1446 AAGAGGATACACTTGGCCCGGTATATGTGTGCAATCATCTCGCATGATTCAGGCT 1505  
Qy 1648 TCTTGGCCAAAGGGAGGCTTGAATCTGTGAGGAAATGAATGACCATGATGTTATATG 1707  
Db 1506 TCTTGGAGATGACGGGTGTCTGTGATGTGAAACAAACGAGTTGCTTCGATTAATGTTACG 1565  
Qy 1708 TTTCTAGAGAAACGACAGGCTATACATCATTAAGAAAGTGTGCTATGAAATGCAT 1767  
Db 1566 TTTCTGTGAGAAAGACCCGATTTTATCATCATTAAGAAAGCTGAGCTATGAAATCCC 1625  
Qy 1768 TGGTCCAGTCTCTGTACTATCAAAATGCTCCATATTTGTTAACTTGAATGTGATC 1827  
Db 1526 TGATACAGATCTCTGGGTTCTATCAATGCTCTTATCTTGAATGTGATGTGATC 1685  
Qy 1828 ACTACATCAACACAGCAAGGCTATTAAGAGAGCAATGTGTTTATGATGACCTTTAC 1887  
Db 1686 ACTACATCAACAAATGCAAAAGCTCTTGAAGAGAGCAATGTGTTTATGATGATCTCATG 1745  
Qy 1888 TAGAAGAAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1947  
Db 1746 CAGAAAGAAATCTGTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1805  
Qy 1948 ACCGATATGCTAACCGGATATGTTCTTTTGTGATCAACATGAAAGTTTGGATGTA 2007  
Db 1806 ATGATATCTCAAAATCGCAATGTTGTGTTCTTTGATATCAATATGAAGTTTGGATGGC 1865  
Qy 2008 TTTAGGTTCAATTTATGTTGTGATCTGATGTGTATTTAGAGGACGAGCATTAATATGCTT 2067  
Db 1866 TACAAAGGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1925  
Qy 2068 ATGATGCCCCCAAAACAAAGAGCCACATCAAGAGATTTGCACTGTGCCCCAAGGT 2127  
Db 1926 TTGATGACCGAAG 1985  
Qy 2128 GCTTTGCTGTGTGCTTGTGCAATAGAGAGCAAAAGAGACTAACAAACCCAAACAG 2187  
Db 1986 GTCTCTATGTTTGTGTTCAAGAAAGATGTAAAGCAAGAGAGAGAGAGAGAGAGAGAG 2045  
Qy 2188 AGAAGAAAGATTATTTTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2247  
Db 2046 AGAAGAGAA-----TAGGAGAGGTCAAGAGAGATCAAGCATTAAGAAATA 2093  
Qy 2248 GTGAATTTGAGAGAGAGTCTCCAGAGCTGAGATGAAAGAGAGAGAGAGAGAGAGAGAG 2307  
Db 2094 TCGAAGAGGCGCGGCTCAAAAGTTCTTAACGTAGAACGTCAACGAGGCAATGCAAA 2153  
Qy 2308 AAAAATTAGAAAGAAATTTGGCCCAATCTTGTGTTTGTGTTACATCCACTTCTGAGA 2367  
Db 2154 TGAAGTTGCAAGAAATATGGGAGTCTCTGATATGTGATGATGATGATGATGATGATGATG 2213  
Qy 2368 ATGTGGAACCTTGAAGAGTGAATGTCCTGCTCTCTTTGAAAGAGACTATACATGCA 2427  
Db 2214 ATGTGGAATGCGCTAGAAAGCAAGCCCGCTGTGCTTAAGAGAGCCATCAAGTCA 2273  
Qy 2428 TTAATTTGTGTTATGAGAGACAGACAGACTGGGGAGAAAGAGATGAGATCTATGAT 2487  
Db 2274 TTAATTTGTGTTATGAGAGATTAAGATGAAATGGGAGAAAGATGAGATGATGATGATGATG 2333

```

QY 2488 CAGTTACAGAGATATTTCTAACTGGTTCAAGATGATGTCATGGTTGGCGTCAATTT 2547
DB 2334 CTGTACACCAAGATATTTCTTACGGGTTCTTAAGATGATTCATGGTTGAGACATGTTT 2393
QY 2548 ACTGCATTAACCTAAACGGGTTGATTCAAAGGTTCTGCACCTCTGAATCTTTGAGATCGTC 2607
DB 2394 ATTGTACACCAAGATTAAGCGGTTTCAAGAGATCACTCCAAATCAATCTTTGAGATCGTC 2453
QY 2608 TTCACCAAGGTGCTCGGTGGGCTCTTGGGCTATAGATCTTCTTCAAGCAATCAATGGCC 2667
DB 2454 TCCATCAAGTCTTGTGATGGGCGCTTGGGCTGGTGAATTTCTTGAATAGGCAATGTGC 2513
QY 2668 CTCTTTGGATGGGATGAGTGGCGGCTGAAATTTTGGAAAAGATTTTCTTACATCAACT 2727
DB 2514 CTATTTGGATGATGATGAGGAGGTGAATAAGCTTGAAGCGGTTGTCTTACATTAATCT 2573
QY 2728 CCAATCGTATATCTTGGACATCAATCCCTCTTGGCTTATAGTAAATGCTCGCATCT 2787
DB 2574 CTGTGGTTTACCGGTGACCTCTTACCGCTCATCGTTTACTGTCTCTCCCTGCAATCT 2633
QY 2788 GTTATTTGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGGCAGCCTGTGGTTCA 2847
DB 2634 GTCTTCTCACTGGAATTAATTCATGCTTCCGAGATTAAGCAACTATGAGATATCTCTTCA 2693
QY 2848 TGTCACTTTTATCTGCATTTTGTGCTACGAGCATCTTGAATAATGAGATGAGTGTGTG 2907
DB 2694 TGGCGCTCTTCTGCTGATTTGCAATACGGGATATCTCGAATGCAATGGGCGAAAGTTG 2753
QY 2908 GAATGTATGATGCTGTGAGGAATGAGCATGTCGGGTCATTTGGAGGTGTGCTCTCAACAC 2967
DB 2754 GGAATCATATATGTTGTGAAGAAACAGATTTTGGGTCAATGAGAGTGTCTTCCGCATC 2813
QY 2968 TCTTTGCTGTGTTCCAGGACTTCTCAAGGTCACTAGCTGTGTTGATACAAAGTTCACCG 3027
DB 2814 TGTGTCTCTCTTCCAAAGTCTCTCAAGGTTCTTGTGTGTGTCGACATTAATCTTACAG 2873
QY 3028 TGAATCAAAAGGTGAGATGATGAGAGTCTCAGAGCTATATACATTTCAAAATGAGCTA 3087
DB 2874 TCACATCAAAAGGAGCTGATGATGAGAGAGTCTCTACCTTAACTCTTCAAAATGAGCT 2933
QY 3088 CCTATGTATGATCTCTCAACCACTTCTCTTATGAACTTCAATGATGTGTGTCGCTGAG 3147
DB 2934 CACTTCTCATCTCTCAATGACTTACTATCAATTAAGGTCAATGAGTCAATAGTCGAG 2993
QY 3148 TTTCAATGATGATCAATTAAGGATATGATGATGAGGCGCCCTCTTGGGAAAGCTATCT 3207
DB 2994 TCTGTATGATGATCAATTAAGGATATGATGATGAGGCGCCCTCTTGGGAAAGCTATCT 3053
QY 3208 TTTGATTTTGGGATGTCATCTTATCTCTTCTCAAAAGGTTTGGTTGAAAGGCAAA 3267
DB 3054 TTGACATTTTGGGATGTCATCTTATCTTACCCGTTCTTAAAGGTTTGGTTGAAAGGCAAG 3113
QY 3268 AAGGACACCAAGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3327
DB 3114 ATGAGATGTCACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3173
QY 3328 TTTGGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3387
DB 3174 TTTGGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3230
QY 3388 GTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3398
DB 3231 GTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3241

```

## RESULT 11

```

AB212754 standard; DNA; 3255 BP.
ID AB212754 standard; DNA; 3255 BP.
AC AB212754;
XX
DT 21-JAN-2003 (first entry)

```

```

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 559.
DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
KW Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN WO20021655-A2.
XX
XX 28-FEB-2002.
PD
XX 24-AUG-2001; 2001WO-US26685.
PF
XX 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-264647P.
PR 22-JUN-2001; 2001US-300111P.
XX
XX (SCRI ) SCRIPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
PI Harper JF, Krepe J, Wang X, Zhu T,
XX WPI: 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX
XX Claim 144; SEQ ID NO 559; 577bp + Sequence Listing; English.
PS
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stresses. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 3255 BP; 884 A; 621 C; 798 G; 952 T; 0 other;

```

```

Query Match 36.4%; Score 1431.8; DB 24; Length 3255;
Best Local Similarity 66.0%; Pred. No. 0; Mismatches 1052; Indels 63; Gaps 4;
Matches 2166; Conservative 0;

```

```

QY 144 ATGAGAGGAGCGCGGCTGTGAGCGGCTCCCAACCGCAAGAGCTGTCATC 203
DB 1 ATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 204 CGCCCGACCGGATCCCGGCGGAGCGCGCGGAGCAAGCGGCGAGGTGCGAG 263
DB 61 AACCGCGATGAGAGGCGCAATACGATCAATCAAGATGAGGCGCAAGATGCA 120
QY 264 ATTTGGCGGAGAGAGTGGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 323
DB 121 ATCTGTGAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 180
QY 324 TGGCGCTTCCCGTCTGCGGAGTGTACGAATTAAGAGCGCGGAGAGGCGAGCGAGAAC 383
DB 181 TGGCGATTTCCGGTTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 384 TGGCGCGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
DB 241 TGTCTCAGATGCAAACTGATACAAAGATTAAGGATGATGATGATGATGATGATGATGAT 300
QY 444 GAGGAGGAGAGCGGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503

```

Db 301 GATGAAGAAGAGAAGATGATGATCTTGAGTATGAGTTTGATCATGGAGTGAACCT 360  
Qy 504 CAGCTGTGGCCGAGTCCATGCTCTAAGCCCAATGAGTACGCGCGTGAAGTAAACCT 563  
Db 361 GAAACATGCGCTGAAGCCGCACTCTTACCCCTTAAACACCGCGTGTGTGATGAT 420  
Qy 564 AATGGCGGCCCAAGCTTCCAGCTCAACCCCAATGTTCCACTCTCCACCAACCGGCA 623  
Db 421 TCAGTCCACCTG-----GCTCTCAATTCCTCTTTGACTTATTTGTAT 465  
Qy 624 ATGTGATGATCATCCACCGAGACGACGCGTGTGCTCTTTTCATGGGTGTGG 683  
Db 466 GAAAGTCTGATATGATTTCTGATGCTCATGCTTTATGTGCTCTTCAACGGGATAT 525  
Qy 684 GGAAGAAGATACATCCCTCTTATGCGGATCCCACTTAAGTGTGACCCAGGCT 743  
Db 526 GGAATCGCGCTCATCTGCAACGGTTTACAGATTTCTTGCACTCCACAGCGCGATCA 585  
Qy 744 ATGACCCATCCAGATCTTGCTCATATGGGATATGATGTTGCTTGAAGAGACGG 803  
Db 586 ATGTTCTCAGAAAGATATGCGGAATATGATATGAAAGTGTGCTTGAAGACCGT 645  
Qy 804 ATGAGAATTTGGAAGCAGACAGAGAGATGACACAGCGGGAATGATGTGTGT 863  
Db 646 ATGGAAGTTGGAAGAGACGACAGCGGAAAGCTTCAAGTCAATTAAGATGAAGAGA 705  
Qy 864 GATGATGT-----GACGATGCTGATCTACCTAATG 896  
Db 706 AACAAATGTCAGATTCCAATGATGACAGAACTAATGATGATCTGACATGCTATGATG 765  
Qy 897 GATGAAGCAAGCAACAACTGTCCAGAAAAATTCACCTTCATCAAGCCAGATTAATCA 956  
Db 766 GATGAAGGAAGCAACCTCTCAAGAAAGCTAATTCATTCGTTCAAGCAAAATTAATCT 825  
Qy 957 TATAGATGATATATATATGCGGCTGTGTTTGGGGTTCCTCTCCACTACCGAGTG 1016  
Db 826 TACAGATTTAATTTCTGTGTGCGCTCGGATTTCTTGTGCTTTCTTTTCAATTAAGAAAT 885  
Qy 1017 ATGATCCCGATGATGATGATGATGCTTTGTGTGCTCATATCTGTATCTGTAAATCTGG 1076  
Db 886 CTCATCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 945  
Qy 1077 TTTCGATGCTTTGATTTCTTGATCAATTTCCAAAGTGTTCCTATTTGAAGAGAGACT 1136  
Db 946 TTTCGATGCTTTGATTTCTTGATCAATTTCCCAATGATGATGATGATGATGATGATG 1005  
Qy 1137 TACCTAGACCGGCTGCACTGAGGTTCAGAGGAAGCCAGCAATCTCAACTGTGCA 1196  
Db 1006 TACCTGATGATGCTCTCTTCAGGTACGAGGAAGGAAGAAACCTCAGGATTTAGCACC 1065  
Qy 1197 ATGATTTCTTTGTCAGTACGCTGATCCCTTAAAGAACCTCTTTGTGTCAACAAT 1256  
Db 1066 GTTGAATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1125  
Qy 1257 ACTGTTCTATCTATCTTTGCGTGAATATCTGTGATTAAGGTTTCTGTATGTTCT 1316  
Db 1126 ACAGTCTTTCCATCTACAGATGATATCTGTGATTAAGGTTGCGGTATGATCA 1185  
Qy 1317 GATGATGCTGCAATGCTAAGTTGAAGATTAATCGAAGCAATCGAATTTGCAAG 1376  
Db 1186 GACGATGCTGCAAGCTTACATTTGAAGCTCTCTGATACAGCTGAGTTGCTGACA 1245  
Qy 1377 AAATGGTCTCTTTCTGCAAAAGCTACATATTTGAACCTCGCGCTCAGAGTGTATTC 1436  
Db 1246 AAATGGTCTCTTTTGTAAAGATTTAATATCGAGCCACGAGCTCTGAGTGTATTT 1305  
Qy 1437 CAACAGAGATGACTACTTTGAAAGACAAAGTGTGACGAACTTTGTTAGGAGAGAGA 1496  
Db 1306 TCTCAGAAATGATTAACCTGAAGAACAAAGTTCACTGCTTTGTGAGGAAAGCTCT 1365  
Qy 1497 GGAATGAAGAAGAGATGAGAAATCAAGGTGAGAAATCAATGCTTATTCGCAAGGC 1556  
Db 1366 GCTATGAAGAAGATTAATGAAGATTTAAAGTGAAGATTAATGCACTGATGCTACTGCA 1425

Qy 1557 CAGAAAGTCTCTGAAGAGATGACAAATGCAAGTGAACCCCTGCTGCAAAACAT 1616  
Db 1426 CAGAAAGTCTCTGAAGAGATGACAAATGCAAGTGAACCCCTGCTGCAAAACAT 1485  
Qy 1617 GTTGTGATCATCTCGAATGATTCAGGCTTCTCTGCGCAAGCGAGGCTTCACTGT 1676  
Db 1486 GTTGTGATCATCTCGAATGATTCAGGCTTCTCTGCGCAAGCGAGGCTTCACTGT 1545  
Qy 1677 GAGGAAATGAATGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1736  
Db 1546 GATGATTAATGATTAACAGCTCTAGTATGATGATGATGATGATGATGATGATGATGAT 1605  
Qy 1737 CATCTAAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1796  
Db 1606 CACCAAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1665  
Qy 1797 GCTCATATTTGTTAACTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1856  
Db 1666 GCTCTTACCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1725  
Qy 1857 GAAGCAATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1916  
Db 1726 GATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1785  
Qy 1917 CCTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1976  
Db 1786 CCGCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1845  
Qy 1977 TTGATATCAATGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2036  
Db 1846 TTGATATTAATGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1905  
Qy 2037 TGTGATTTAAGAGCAGCATTAATGATGATGATGATGATGATGATGATGATGATGATG 2096  
Db 1906 TGTGATTTAAGAGCAGCATTAATGATGATGATGATGATGATGATGATGATGATGATG 1965  
Qy 2097 TCAGGACTTCAACTGCTGCGCAAGTGTGCTTTGCTGTGCTGCTTTGGCAATAG 2156  
Db 1966 GGCAGAAACCTGTAATGCTGTGCTCAAAATGAGTGTGCTTTGCTGTGCTTTGAGAAAG 2025  
Qy 2157 AAGCAAGAGAGATCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2216  
Db 2026 AGTAAAG 2073  
Qy 2217 GAAG 2276  
Db 2074 -----AGCAGATTCAGCCCTAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAG 2127  
Qy 2277 GAGATGAAG 2336  
Db 2128 AATGTTGAG 2187  
Qy 2337 TCTGTTTGTATCACTCAACTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2396  
Db 2188 CCGGTTTGTGCTCTGCTGTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2247  
Qy 2397 GCTTCTCTTTGAAG 2456  
Db 2248 GATGTTTGTATGAAG 2307  
Qy 2457 TGGGAAAGAGAGATGCTGATCTATGATCACTTACAGAGAGATTTCTAATGCTGTTTC 2516  
Db 2308 TGGGAAAGAGAGATGCTGATCTATGATCACTTACAGAGAGATTTCTAATGCTGTTTC 2367  
Qy 2517 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2576  
Db 2368 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2427  
Qy 2577 GGTCTGCACTGCAATCTTCAAGATGCTTCAACAGAGTGTGCTGCTGCTGCTGCTGCTG 2636  
Db 2428 GATGCTGCTCTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGATG 2487

```

QY 2637 TCTATTGAGATCTTCTTCACGATCATTCGCCCTCTTGGTATGGGATGTGGCGGTCTG 2696
DB 2488 TCTGTAGAGATTTCTTGAGCAGACATTCGGAATATGGTATGTGTGTGTGTGTTA 2547
QY 2697 AAATTTTGGAAAGATTTCTTACATCAATCCATCGGTATGCTTGGACATCTATGCC 2756
DB 2548 AATATGTTGGAAGATTTCTTACATCAATCTGTGCTTATCTTATCTTGGACTTACTTCA 2607
QY 2757 CTCTGGCTTACTGTACATTCGCTGCACTGTGTTATTTAGACAGGAATTTATCACTCA 2816
DB 2608 TTGATCGTCTATTGTTCTCTCCCGGGTGTGTTTACTCAGAGAAATTCATCGCCCT 2667
QY 2817 GAGCTGAATATATGTTCCAGCCTGTGTTCAATGTACATTTTATCTGCACTTTTGTACG 2876
DB 2668 GAGATAGACACTACGACGATATCTCTCATGCTCATGTTCATATCCATACAGAACT 2727
QY 2877 AGCATCTGAATATGATGAGATGGTGTGTAATTTAGATGACGTGGAGGAATAGCAG 2936
DB 2728 GGAATCTCTGAATATGATGAGATGGGAGGTGTGGATGTGATGTTGGGAAACGAGCAG 2787
QY 2937 TTCTGGGCTATGAGAGTGTCTTCACACACTCTTGTGCTGTGTCCAGGACCTTCTCAG 2996
DB 2788 TTTTGGGTATGAGAGGGGCTCTCTGCAATATTTGCTGTGTTCAAGTTTGTCTCAA 2847
QY 2997 GTCATAGCTGTGTGTATCAAGCTTCCAGCTGACATCAAGGGTGGAGATGATGAGAG 3056
DB 2848 GTTCTAGCCGAGATTAAACAGAAATTCACAGTCACTTCAAAAGCAGACATGAGAGCT 2907
QY 3057 TTCTGAGCTATATATCATTTAAATGAGACTTATGATATCCCTTCCACCTTGTCT 3116
DB 2908 TTCTGAGCTTATATCTTCAAGTGAACATTTGTATATCTTCCAGCAACACTTCTG 2967
QY 3117 CTATTGAACTTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3176
DB 2968 ATCATTAATATCATTTGAGATTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3027
QY 3177 TCATGAGGACCCCTCTTGTGGAAGCTATTTCTTGTGTGTGTGTGTGTGTGTGTGTGT 3236
DB 3028 TCATGAGGACCTCTCTTGTGGAAGCTATTTCTTGTGTGTGTGTGTGTGTGTGTGTGT 3087
QY 3237 CCCTTCTGAAAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3296
DB 3088 CCAATTCCTCAAGGAAATGCTTGTGGAAGCAAGAAATGCTTACGATTAATGTGTGTGTGT 3147
QY 3297 TCCATTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3356
DB 3148 TCAATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3207
QY 3357 AAGGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3397
DB 3208 A--AAGGGGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3245

RESULT 12
AAV06566
ID AAV06566 standard; cDNA; 3828 BP.
XX
AC AAV06566;
XX
DT 06-JUN-1998 (first entry)
XX
DE Arabidopsis cellulose biosynthetic gene clone Ath-A.
XX
KM Cellulose; cellulose synthase; RSW1 gene; beta-1,4-glucan;
XX
OS transgenic plant; ss.
XX
XX Arabidopsis thaliana var. Columbia.
XX
XX
XX Key Location/Qualifiers
XX CDS 239..3490
XX FT /*tag= a
XX
PN MO9800549-A1.

```

```

XX 08-JAN-1998.
PD
XX
XX 24-JUN-1997; 97WO-AU00402.
PF
XX
XX 27-JUN-1996; 96AU-0000699.
PR
XX
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (AUSU ) UNIV AUSTRALIAN NAT.
PI
XX
XX Arioli A, Betzner AS, Peng L, Williamson RE;
DR
XX WPI; 1998-086974/08.
DR
XX P-PSDB; AAM33818.
PT
XX
XX DNA encoding cellulose biosynthetic enzyme - useful for manipulation
PS of cellulose and beta-1,4-glucan
PS
XX Claim 12; Page 123-130; 207pp; English.
XX
XX cDNA clone Ath-A was isolated from an Arabidopsis thaliana cDNA
XX library using probes produced by PCR with primers (see AAT9632-34)
XX based on cellulose synthase RSW1 genomic clone 23H2 (see AAV06563)
XX and EST clone AAT20782 (see AAV06562). It is closely related to RSW1
XX cDNA (see AAV06565) and to partial genomic clone 12C4 (see AAV06564),
XX and encodes a 1084-polyprotein (see AAM33818). Claimed nucleic acid
XX molecules (see AAV06562-69) coding for claimed polypeptides (see
XX AAM33816-20 and AAM46202) involved in cellulose biosynthesis can be
XX used to manipulate the cellulose and/or beta-glucan content of
XX transgenic plants. Expression of nucleic acids in the sense
XX orientation increases the level of cellulose and reduces the level
XX of non-crystalline beta-1,4-glucan and starch, providing plants
XX with modified strength and/or shape and/or fibre properties, or
XX having increased resistance to stresses or pests. Antisense,
XX ribozyme or co-suppression molecules can be used to reduce the
XX cellulose content of a transgenic plant, e.g. to improve
XX digestibility or to alter carbon partitioning such that increased
XX carbon is available for growth, rather than deposited as cellulose.
XX
XX Sequence 3828 BP; 1056 A; 721 C; 912 G; 1139 T; 0 other;
XX
XX Query Match 36 3%; Score 1429.4; DB 19; Length 3828;
XX
XX Match Local Similarity 65.8%; Pred. No. 0; Mismatches 1066; Indels 63; Gaps 4;
XX
XX Matches 2172; Conservative 0;
XX
QY 124 GAATGAGAGGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
DB 219 GATTGGGTGTAGAGACATCATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 278
QY 184 GCACGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
DB 279 GAAACGAAATTCGTTCTCATTAACGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
QY 244 AGAACGGGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
DB 339 TGAGTGGGCAACATGATTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 398
QY 304 CCTTGTGAGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
DB 399 TCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 458
QY 364 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
DB 459 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
QY 424 GCGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
DB 519 GTCCAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 578
QY 484 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
DB 579 TTGATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 638

```

Oy	544	ACGGCCGGGAGAGTGAACCCCTAATGGCCGGCCACAAGCTTTCAGACTCAACCCCAATGTC	603
Db	639	CCGGTCCGGTGGAGATTGGATTCACTCACCTGCCTC-----CAGATTC	683
Oy	604	CACCTCTCAACAACGGGCAATGGTGATGACATCCCAACGGAGACAGCCGCTGGTGC	663
Db	684	CTCTTTTGACTTATTGTGATGAAAGTGTGATATGTATTCTGAATGCTCATCTCTTATCG	743
Oy	664	CTTCTTTTCATGGGTGGGGGAAAAGAGATACATCCCTTCCTTATGCGGATCCAGCT	723
Db	744	TGCCTCTCTCAACGGGATATGGGAATGCGCTATATCTGACACCGTTTACAGATTCTTCTG	803
Oy	724	TACCTGTGCAACCCAGGCTCTATGAGACCATCCAGAGATCTTGCTGCATATGGGATGGTA	783
Db	804	CACCTTCCACAGCCGAGATCAATGGTCTCTCAGAAAGATATGGCGGAATATGGTATGGA	863
Oy	784	GTGTGTCTTGGAAAGGAACGATGTGAGATTGGAAAGCAGACAAGAGAGATGACCGA	843
Db	864	GTGTTGCTTGGAAAGGACCTTATGGAAGTTTGGAAAGACGACACAGGCCAATAAGCTTCAAG	923
Oy	844	CGGGGAATGATGATGGTGGTATGATGCT-----GACG	876
Db	924	TCATTTAACATGGAAGAGAAACAATGATGAGGTTCCATATGATGACGACGAATGATG	983
Oy	877	ATGCTGATCTACCACTAATGATGAGCAAGACAACAACCTGTCCAGAAAAATTCACATTC	936
Db	984	ATCTGATCATCTTATGATGATGAGGAAAGACCACTCTCAAGAAAGCTACCTATTTC	1043
Oy	937	CATCAGGCACATTAAATCCATATAGGATGATTATCATTTCCGCTGGTGGTTGGGGT	996
Db	1044	GTTCAAGAGAAATAAATCTTACAGAGATGTTAATCTGTGTCCTCCGAAATCTTGCTC	1103
Oy	997	TCTTCTTCCACATACGAGATGATGATCATCCGGTGAAATGATGCAATTTGCTGTGCATAT	1056
Db	1104	TTTTCCTTCAATATAGATATCTCCATCCAGTCAAGATGCAATATGATTAATGGTTAAAGT	1163
Oy	1057	CTGTATCTGTGAATCTGTTTGGCATGTCTTGATCTTGATCAATTTCCCAAAATGTGT	1116
Db	1164	CAGTTATATGGAAATATGTTTGGCAGTGTCTTGATTTCTTGATCAATTTCCCAAAATGTGT	1223
Oy	1117	TCCCATATGAGAGAGACTTAACCTACAGCCGGCTGTCACTAGGTTGCAACAAGAAAGGCC	1176
Db	1224	ATCTCATAGAACGTGAACAATACCTCGATAGACTCTCTCTCAGGTAACAAGAAAGGAA	1283
Oy	1177	AGCCATCTCAACTTGCTCCAATATGATTTCTTTGTCAAGACGGTGTATCCCTTAAAGAAC	1236
Db	1284	AACCGTCAGAGATTAGCACTGTGATGTTTTTTGTTAGTACAGTGAATCCGTTGAAGAGC	1343
Oy	1237	CTCCTTGGTCAACAATACTGTTCTATCTATCTTTTGGGTGATTAATCCTGTGTATA	1296
Db	1344	CCCCCTTGAATACAGCAACAACAGTTCTTTCATCTTAGCAGTGAATATCTGTGATATA	1403
Oy	1287	AGGTTTCTTGCTATGTTTCTGATATGATGTGTCGAATGCTTAACGTTGAACATTAATCTG	1356
Db	1404	AGTTGCGGTATATGATATCAACAATGTGTGACGCTATCTTACATTTGAACCTCTCTGTG	1463
Oy	1357	AAACATCTGAATTTGCAAGAAATGGTGTCTTCTTGTGCAAAACGGTACAATATTGAACCTC	1416
Db	1464	ATACAGCTGATTTTGTCTACAAATGGTGTCTTTTGTGAAGATTTAATATCAGACAC	1523
Oy	1417	GGCGTCCAGATGTACTTCCAAACGAAGATAGACTACTTGAAGAACAAGGTGGCAGCA	1476
Db	1524	GAGCTCCGAGTGGATATTTTCTCAGAGATGAGATTACTGTGAAGAACAAAGTTCAATCCTG	1583
Oy	1477	ACTTTGTTAGGAGAGAGAGCAATGAGAGAGAGTATGAGGAATTCAAAGTGAAATCA	1536
Db	1584	CTTTTGTACAGGAAGTGTGTGTATGAAGAGAGATTATGAAGAGTTTAAATGTGAAGATTA	1643
Oy	1537	ATGCTTATGTTGGCCAAAGCCCAAGAAAGTTTCCGAAGAGATGACCAATGCAAGATGGA	1596
Db	1644	ATGACACTGTTGCTACTGACAGAAAGTGTCTGAGGAACGTTTGAATATATGCAAGATGGA	1703
Oy	1597	CCCCCTGGCCTGGAACAATGTTCTGTATCATCTCGAATGATTAACAGTCTTCTTGACC	1656

Db	1704	CTCCTTGCCCTGGAAAAACAAGTCCGTGACCAATCCTGGAAATGATTCAGGTGTTCTTGGGCT	1763
Oy	1657	AAAGCGAGGCCCTTGACTGTGAGGGAATGAACCTGCCAGATTGGTTAATGTTCTAGAG	1716
Db	1764	ATAGTGGAGTTGCGATACGGAATGGTATGAGTTACCAACGCTCTAGTGAATGTTCTGTG	1823
Oy	1717	AGAAACCAACGAGCTATPACCATATPAAAGAAAGCTGGCTATGATGCAATGGTCCGAG	1776
Db	1824	AGAAAGCGGCTCGAGATTGATACCAACAAGAAAGCTGAGCTATTAATCTCTGTATCCGAG	1883
Oy	1777	TCTCGCGCTATCTPAAACAAATGCTCCATATTTGTTAACTTGGAATTGGAATGACATACCA	1836
Db	1884	TCTCTGCTGTTCTATCAAAAGCTCTTACCTCTTAAATGTGATTTGATCACTACATCA	1943
Oy	1837	ACAAACAGACAGCTATPAAAGGAACAAATGTGTTTATATGATGACCCCTTACTAGAAAGA	1896
Db	1944	ACAAACAGAAAGCAATPAGAAATCTATATGTGTTCAATGATGACCCGCAATCGGGAAGA	2003
Oy	1887	AGGTTTGCTATATACAGTTCCCTCAAAAGATTGATGGGATGATGCCATGACCGATATG	1956
Db	2004	AAAGTTTGTATATGTTCAAGTTTCCGACAGATTGTATGGGATGATPAGCATATATGATACT	2066
Oy	1957	CTAACCGGAATGTGTCTTTTATATPACAAATGAAGAGTTTGAATGATATCAGGCTC	2016
Db	2064	CAAAACCGTAAAGTTGTGTTCTTTATATPAAATGAAGAAAGCTTGATGGGATPACAAAGAC	2123
Oy	2017	CAATTTATGTGTGCTACTGATGTGTATTTAGAAAGCAGGCACTPATTATGTTATATGATGCC	2076
Db	2124	CGATATATATGTGGAGACAGGTGTGTGTTTGAAGAAACAGGCTCTTATATGTTTATGATCAC	2183
Oy	2077	CCAAACCAAGAACCCACCATCAAGACTTGCAACTGTGAGCCAGTGATGCTTTTGCT	2136
Db	2184	CAAGAAAGAAAGAAACCAACAGGCAAAACCTGTAAGTGTGAGCTPAAATGGTGTGTTGTT	2243
Oy	2137	GTTGCTGCTTGGCAATPAGGAAGCAAGAAAGAACTAACCAACCCAAACAGAGAAGAAA	2196
Db	2244	GTTGTGGGTGAGAAAGAAAGTAAACAGAAAGCAAGATPAGAAACCTPAAAGCTPAAAG	2303
Oy	2197	AGTATTATATTTTCAAGAAAGAAAGAACCAATCCCCGTCATATGCTCTTGAGTAAATTTG	2256
Db	2304	AGACTTCA-----AAGCAGATTTCAATGCTGCTAGAGAAATGCTGACAGAG	2345
Oy	2257	ACGAAGCTGCTCCAGAGAGCTGAGATGAAAGAGCCGGTATTTGTAATCAACAAAAATTAG	2316
Db	2346	GTTGTTATGTCCTCCAGTGTCAAAATGTGAGAGAGATCTGAGCAACCAATTGAATTGG	2405
Oy	2317	AAAAAAATTTGGCCAATCTTCTGTTTGTATCATCAACTTCTCGAAGATGTGGA	2378
Db	2406	AGAAAGATTGGACAATCTCCGGTTTTCGTGCTCTGCTGTTCTACAGAACGGTGAAG	2465
Oy	2377	CCTTGAAGAGTGAAGTCTGCTTCTCTTTGAAAGAAAGCTATCATGTCAATTAATGTTGTG	2436
Db	2466	TTCCCCGTAACGCAAGCCCGCAGATGTTGTTAAGAGAAAGCAATTCAGATATTAAGCTCG	2523
Oy	2437	GTTATGAAGACAGACAGACTGGGGAAGAGATGGCTGTGATCTATGATCAAGTTACAG	2496
Db	2526	GATACCAAGATPAAACCGAATGGGGAAGAGATCGGGTGAATTTATGATCGGGAATG	2585
Oy	2497	AAAGATTTCTAACTGTGTTTCAAGATGCAATGTCAATGTTGGCCGCTCAATTTACTGTCAATAC	2556
Db	2586	AAAGATATCTGACGGGTTTCAAGATGCAATGCCATGATGAGATCTGTATCACTGTATGC	2645
Oy	2557	CTAAACGGGTTGCAATCAAAAGGTTCTGCACTCGAATCTTTCAGATGTGCTTACCAAG	2616
Db	2646	CTTAAGCGTGACGCTTTTAAAGGATGTGCTCTTAATTAACATTCAGATGTCTTCAACAG	2705
Oy	2617	TGCTTCGGTGGGCTCTTGGGCTATATGAGATCTTCTCAGCAATCAATGACCCTCTTGGT	2676
Db	2706	TTCTATCGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAAGCAACATTTGTCCGATATGGT	2765
Oy	2677	ATGGGATATGTGGCGGTCTGAATTTTGTGAAAGATTTTCTTACATCACTCATCGTGT	2736



Db 666 TCATACACGAGCAACCAATTGGCTTGCACACCTCGCAATGCTGAGACGCTGTCATCA 725  
Qy 708 TATGCGATCCCAAGTTACTCTGTGCAACCCAGCTGTATGACCCATCCAGATCTTGCT 767  
Db 726 AGCTCCGAGCCGGG-AGTGAAGATCATCATGATGCCAAGGAAATATGCT 779  
Qy 768 GCATATGGGTATGATGCTGCTTGAAGAAACGATGAGAAATTTGACACAGACAA 827  
Db 780 TCTTATGGCTTTGGAGCGTGTCTTGAAGAGACAGCGCATGGTTATTAATCGAAGAA 839  
Qy 828 GAGAGATGACACAGACCGGGAATGATGCTG- 862  
Db 840 AACCAATAGGCGCAGTTGATATAGCGAAGGAGATCAATATATGCGGGGTTTGA 899  
Qy 863 -----TGATGATGTCAGATGCTGCTTACCACTATATGATGAGACAA 911  
Db 900 CCAATAGACCTGAAGATTATATGATCCCATATGACCAATGACGATGAAAGCAAGGAG 959  
Qy 912 CAACGTCCAGAAATTCACCTTCATCAAGCCGATTAATCCATATAGATGATATTC 971  
Db 960 CCACGTCCCAAAAGTCCCAATTCCTCAAGCAAAATTAATCCATACCGATGGTCAAT 1019  
Qy 972 ATTATCCGCTTGTGCTTGGGGTCTTCTTCCACTACGAGATGATCCGGTCAAT 1031  
Db 1020 GTATTTGACATGATGCTGGTATTTTCTCCGCTATCGTCTCTGAATCCAGTAG 1079  
Qy 1032 GATGCAATTTGCTTGTGCTCATATCTGTATCTGTGAATCTGTGTTTGCATGCTTGG 1091  
Db 1080 AATGATATGGGCTCTGGGCCACTTCTATCTGTTGTGAATCTGGTTTCCCTTGCTATGG 1139  
Qy 1092 ATTCTGATCAATTCACCAAGTGTTCCTTATGAGAGAGACTTACTAGACCGGCTG 1151  
Db 1140 ATCTCTGATCAAGTTCCCAAGGTGTGCTTCACTGATCGTGAAGGATATCTGATGACTG 1199  
Qy 1152 TCACGTAGGTTGCAACAGAAAGCCAGGCACTCTCAATGCTTCCAAATGATTTCTTGTG 1211  
Db 1200 TCATTAAGTACGAAAGAGAGGCAACATCAATGCTTGCACCTGTTGACCTCTTGTG 1259  
Qy 1212 AGTACGGTTGATCCCTTAAAGAACCTCTTGTGTCACAAACAAATCTGTCTATATC 1271  
Db 1260 AGTACTGATATCCACTAGAGAGGCTCTTGTGTTACTGCAATATAGTATATCAATC 1319  
Qy 1272 CTTTCGGTGAATATCTCTGTGATTAAGTTCTTCTATGTTTCTGATGATGCTGCA 1331  
Db 1320 CTTTCAGTAGACTCCCTGTGACATATGTCCTGTATGTCCTGATGAGAGGCGCTG 1379  
Qy 1332 ATGCTAACGTTGAAGCATTTATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTC 1391  
Db 1380 ATGCTTACTTTGAATCTCTCTCTGAGACCTCAGAAATTTGCCAAGAAATGGGTACATTC 1439  
Qy 1392 TGCAAAACGGTACAATATTTGAACCTTCGCGCTCCAGAGTGTACTTCCACAGAGATGAC 1451  
Db 1440 TGCAAGAAATTCGACATTGAGCGCTCGCGCTCCGAAATCTATTTCTCTCAAGAAATTTAC 1499  
Qy 1452 TACTTGAAGACAGAGTGCAGCAAACTTTGTAGAGAGAGAGACATGAAGAGAG 1511  
Db 1500 TATCTGAAGGACAAATTTCAACCCACTTTGTCAAGAGCGCCGTCATGAAGAGAA 1559  
Qy 1512 TATGAGAAATTCAGAGTGAATCAATGCTTATGTTGCAAGGCCCAAGAAATTTCTGAA 1571  
Db 1560 TATGAGAAATTCAGAGTGAATCAATGCTTATGTTGCAAGGCCCTTAAAGTGTCCAG 1619  
Qy 1572 GAAGATGAGACATGACAGATGAAACCCCTGCGCTCGAAGAAATGTTCTGATCATCT 1631  
Db 1620 GAAGATGAGACATGACAGATGAAACCGCTTGGCTGTGTAATTAATACCGTGAACATCT 1679  
Qy 1632 GGAATGATTCAGGTTCTCTTGGCCAAAGCGAGGCTTGTACTGTAGAGGAAATGAAGT 1691  
Db 1680 GGTATGATTCAGGTTCTCTTGGGTCACAGTGGCGCTCTGATACAGAAAGGAAATGAAGT 1739  
Qy 1692 CCAAGATTTGTTTATGTTTCTAGAGAAAGCAACGAGCTATTAACCATCATAGAAAGCT 1751

Db 1740 CCTCGGCTAGTATATGTTTCTCGTAGAAGAGACCTGTTTCCAGCATCACAAGAGCC 1799  
Qy 1752 GGTGCTATGATGATGATGCTCCAGTCTCTGCTGTACTACAAATGCTCCATATTTGTA 1811  
Db 1800 GGTGCTATGATGATGATGATGCTCCAGTCTCTGCTGTACTACAAATGCTCCATATTT 1859  
Qy 1812 AACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1871  
Db 1860 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919  
Qy 1872 ATGATGACCTTTTACTAGAGAAAGAGTTTGTCTATGTACATGCTTCCAAAGATTTGAT 1931  
Db 1920 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1979  
Qy 1932 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1991  
Db 1980 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
Qy 1992 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2051  
Db 2040 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099  
Qy 2052 CAGGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2111  
Db 2100 CAGGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2139  
Qy 2112 TGCTGCGCCCAAGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2171  
Db 2140 -----GTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2195  
Qy 2172 ACCAAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAAAGCAATCC 2231  
Db 2196 CCGAAGAAATTAATGACAAAGAAAG-----TGAGAAATCCAGCT 2237  
Qy 2232 CTTGCAATATGCTCTTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2291  
Db 2238 CCGGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297  
Qy 2292 GGTATGTAATCAACAAATTAAGAAAGAAATTTGGCCAAATCTTGTGTTTGTACA 2351  
Db 2298 TGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2357  
Qy 2352 TCCACACTCTCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2411  
Db 2358 TCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2417  
Qy 2412 GAAGCTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2471  
Db 2418 GAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2477  
Qy 2472 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2531  
Db 2478 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2537  
Qy 2532 GGTGCGGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2591  
Db 2538 GGTGCGGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2597  
Qy 2592 AATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2651  
Db 2598 AATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2657  
Qy 2652 TTAGCAATCATGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2711  
Db 2658 ATGAGCAGACATGCTCCCAATCTGGATGATGATGATGATGATGATGATGATGATGATGAT 2717  
Qy 2712 TTTTCTCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2771  
Db 2718 TTTGCTATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2777  
Qy 2772 ACATGCTGCAATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2831  
Db 2778 ACATGCTGCAATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2837



QY		2832	GCACGCCCTGGTTCATGCACTTTTATTCGCAATTTTGTGCTACAGACATCCTAGAAATG	2831
Db		2838	GCAGGTCAATTTTATATAGCTCTTTTTCATCTCAATTTTGTGCGACATGTATCTGAAATG	2897
QY		2892	AGATGAGATGGTGTGGATGATGATGACGTGAGGAATGACAGTTCGTGGGTCTATTGSA	2951
Db		2898	AGGTGAGATGGAGGTAGCATTTGAAATAATGTGTGGCAATATGAACATTCGTGGGTATTGGA	2955
QY		2952	GGTGTGTCCTCAACACTCTTTTGCTGTGTTCAGGAGCTTCTCAAGGTCAATGCTGTGTT	3011
Db		2958	GGGGTTTCTGCACTTTTGTTCAGATTATTCAAGGCTGTGCTCAAGGTACTGGCAGGCATT	3017
QY		3012	GATACAACTTCACCGGTACATCAAAAGGTGAGATGATGAGAGATTCAGAGCTATAT	3071
Db		3018	GATACAAATTTTCACAGTCACTGCGCAAGGATCAGATGACGTGATTTGGGGAACGTGAT	3077
QY		3072	ACATTCAATGGACACTATTATGATPACCTCCTACACCTGCTCTATTTGAATTCATT	3133
Db		3078	GCATTCAANTGSAACACACTCCTCATTCCTCTTCAACCCGTCTGTGATCAACTGTGTG	3137
QY		3132	GGTGTGTGCTGCGCGTTTCAATGCGATCAATTAAGATATGAGTCAATGGGCCCCCTC	3191
Db		3138	GGGGGTGTGTGTGGGTAGCAGATCAATCAACATGATTCAGTCAATGGGGTCTCTC	3197
QY		3192	TTTGGGAGCATATTCCTTTGGCATTTTGGGGATGTCATCTTATCCCTTCTCAAAAGT	3253
Db		3258	CTCATGGGCGACGACAACCGAACAACCCACACGTCGTGATTATTTGTCATTTCTGTGGCA	3317
QY		3312	TCATCTTCTGCGCTCCTTTGGGTTGCGATTGATCCTTTCCTTGGCAAGATGATGCTCG	3371
Db		3318	TCTGTTTCTCTCTTTTCTGCGGTAAAGAAATTGATCCCTTTGAGTAAAGGTTAAAGGCCCA	3377
QY		3372	CTTCTGAGAGGTGAGTTTGATTCG	3398
Db		3378	GATACTAAACATGTGGCATCACTGC	3404
RESULT 14				
AAZ99500				
ID	AAZ99500	standard; DNA; 3725 BP.		
XX	AC	AAZ99500;		
XX	DT	03-JUL-2000 (first entry)		
XX	DE	DNA encoding a maize cellulose synthase.		
XX	KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;		
XX	transgenic plant; plant breeding marker; ss.			
OS	Zea mays.			
XX	Key	Location/Qualifiers		
XX	FT	179..3398		
XX	FT	/tag= a		
XX	FT	/product= "cellulose synthase"		
XX	FT	/note= "no termination codon given"		
XX	PN	WO200009706-A2.		
XX	PD	24-FEB-2000.		
XX	PF	16-AUG-1999; 99WO-US18760.		
XX	PR	17-AUG-1998; 98US-0096822.		
XX	PA	(PION-) PIONEER HI-BRED INT INC.		
XX				

Pt Dhuga KS, Helentjaris TG, Bowen BA, Wang X, 1997. WPI: 2000-224343/19. DR P-PsDB; AAY84110.

New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein -

Claim 1; Page 105-110; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

Query Match 35.0%; Score 1378.4; DB 21; Length 3725;  
Best Local Similarity 66.4%; Freq. No. 0;  
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9.

Y CGCGGAGTGGAGGAGGAAGCATGTGAGGCGCGGAGTGATGGCCGACTGCCA 178  
D CGGTGTGTGTGAGGAGACCGCGAGATGGCGGCACAAGGAGATGTGTGCAAGCTTCTCA 213  
Y CAACCGCAAGAGCTGTGTATTCGCGCAGCGCA-----TCCGAGCGCAAGCC 232  
D CAACCGCAAGAGCTGTGTATTCGCGCAGCGCA-----TCCGAGCGCAAGCC 273  
Y GCCGCGGAGAGCGAAACGGGAGGTGTGCAATTGGCGCGCACAGACGTGCGCTTCCCCC 292  
D CACGGAAGTGTGCAATTGGGCAAGTGTGCAATTGTGTGCGCAACTGTGGGTTCGATGACG 333  
Y CGCGCGGAGACCCTTCTGTGTGAGGTGCAAGATGTGCGCTTCCCCTGTGCGGAGCTGCTA 352  
D CACGTGTGTATGTCTTTGTGTGTGCTGCAATGATGTGTGCTTCCCTGTGTGCGCGCTTGTCTA 393  
Y CGAATACAGAGCGCGGGAGGGGACGACAGACTGCCCCCAAGTGCAGAGACTCGAATACAGCG 412  
D TGAGTAGCAGAGCGCAAGGAGGAAACCAATGTGTGCGCTTCAAGTGCAGAGACTAGATACAAAG 453  
Y CCTCAAGAGCTGCGCAAGCTGTGACCGGAGAGAGAGAGCGCGCGTGCATGACTGTGA 472  
D ACAGAAAGTATGACCTCGAGTTTCATGTGTGATGATGAGAGAGAAATGTGTGATGACTGTGA 513  
Y CAACGAGTTCAACTGGGACGGCGCATGACTGTGTGAGCCGAGTGCATGTCTTAACG 532  
D CAATGAATTCACATAA-----GCAAGGCAATGGGAGG 549  
Y CCACAATGACTACGGCGCGTGTGAGGTGACCTTAATGGCGCGCAAGCTTTCAGCTCA 592  
D CCGAGAGTGGAGCTTCAAGAGATGAGCGCTGATGTGTTCATGTGTGCTGCGCATGACCC 609  
Y CCCCAAT---GTTCCACTCTCACCAACGGGCAATGTGTGATGATGCCACCGGAGCA 649

Db 610 ACACATCGAATTCCACCGCTTACAGTGAACAAGATATCTGAGAGATCCCTGATGC 669  
QY 650 GCAAGCGGTGGTGGCTTCTTTCATGGTGGTGGGGAAGAAGATATACCCCTCCCTTA 709  
Db 670 ATCCCTGACCGTCAATCTATCCGAGT-----CCAACTCGAGCTA 711  
QY 710 TGGGATCCAGCTTACCTGTGCAACCGAGTCTATGAGCCATCCAGATCTTGTGTC 769  
Db 712 TGTGATCCAAACGCTTCCAGTTCCTGTGAGATTTGTGACCCCTCGAAGACTTAATTC 771  
QY 770 ATATGGTATGTGTGTGTCTTGGAAAGAACGATGAGAAATTGAAAGCAGAGACAAGA 829  
Db 772 CTATGGGCTTAATAGTGTGATCGAAGGAAGAGATTGAGACTGAGAGGCTTAACAGGA 831  
QY 830 GAGATGACCAAGACGGGGAATGATGTGTGTGTATGTGTG----- 873  
Db 832 CAAAAATATGTGGCAAGTGAATTAATATCCAAAGCTTAGAGAGACATGAGAGGGAGC 891  
QY 874 -----ACGATGTGATCTACCACTAATGAGATGAGCAACAACATGTCCAGGA 925  
Db 892 TGGCTCAAAATGAGAAAGATATGCAAAATGTTGATATGACCGCTAACCTTGAAGCCGAT 951  
QY 926 AATTCACATTCATCAAGCAGATTAATTCATATAGATGATTAATCATTAATTCGGCTTGT 985  
Db 952 TGTGCAATTCCTCAACAGCTCAACCTTAACGGATAGTAATCATTCCTCGTCTTAT 1011  
QY 986 GGTTTGGGGTCTTCTTCCATACCGAGTATGATCCGGTGAATGATGATTTGCTTT 1045  
Db 1012 CATCTGTGCTTCTTCTTCAATATCTGATCAAGTATCAAGTCACTTAATGTGATTA 1071  
QY 1046 GTGGCTCATATCTGTTATCTGTGAATCTGTTTGCATGTCTTGATTTCTTGATCAAT 1105  
Db 1072 GTGGCTAGTATCTGTATCTGTGAGGTCTGGTTGCTGTGCTTCTGATCAATCACTT 1131  
QY 1106 CCCAAAGTGTCCCTATTTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTGCA 1165  
Db 1132 CCCAAATATGATCAATCAACCGTGAAGACATCTCGACAGGCTTGATGAGGTATGA 1191  
QY 1166 CAAAGAGCCAGCATCTCAACTGTCTCAATTTCTTTGTGAGTACGGTGTATCC 1225  
Db 1192 TAGAAGGAGAGCATCAAGCTGGCTCCCATGTATGCTTGTCACTAGTGAATCC 1251  
QY 1226 CTTAAGGAACCTCCTTGTGTGACAAATACTGTCTTCTATCTCTTGTGCTGATTA 1285  
Db 1252 ATTTGAAGAACTTCACTGATTAACAGCAACCTGTGTCTTCTTGTGCTGATTA 1311  
QY 1286 TCTGTGTATAGTCTTCTGTATGTGTGTGATGATGATGATGATGATGATGATGAT 1345  
Db 1312 CCTGTGTACAAAGTGTATGCTATGCTTGTGATGATGATGATGATGATGATGAT 1371  
QY 1346 AGCATTAATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACCGTACA 1405  
Db 1372 GTCTCTCTGAAACCTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTGTAAGAACACA 1431  
QY 1406 TATTGAACCTGCGCTTCCAGAGTGTATCTTCAAGAAAGATAGCTACTTGAAGACA 1465  
Db 1432 TATTGAACCAAGAGCTCAGAAATTTTACTTGTCTCAAAAATATGATTAACCTGAAGACA 1491  
QY 1466 GGTGACAGCAAACTTTTGAAGAGAGAGCAATGAAAGAGATGATGAGATTTCA 1525  
Db 1492 AATTCAACCTTCAATTTTGTAGAAAGAGAGCAATGAAAGAGATGAAAGATTTCA 1551  
QY 1526 GGTGGAATCAATGCTTATGTTGCCAAAGCCAGAAAGTTCTTGAAGAGATGAGCAAT 1585  
Db 1552 AATAAGATCAATGCTTGTGTGCTTGAAGCAGAAAGTGTCTGAAGAGGCTGAGCAT 1611  
QY 1586 GCAATATGGAACCCCTGCGCTTGAAGCAATGTTGTGATCATCTTGAAGATATTCAGGT 1645  
Db 1612 GGTATATGGAACCTTGTGCTTGGGAATTAACCTTGAAGCAATCTTGAAGATTCAGGT 1671  
QY 1646 CTTCCTTGGCAAGCGAGGCTTGAAGTGAAGGAAATGAAGTGCACGATGATGATTTA 1705  
Db 1672 GTTCTTGGGACAGTGTGGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1731

QY 1706 TGTTCCTAGAGAGAAAGCAGCAGCTAATTAACCATCATTAAGAAAGCTGTGATGAATGC 1765  
Db 1732 TGTCTCTGTGAAAGAGACCGAGCTTTCAGATCAACAAGAGCTGTGTGATGAATGC 1791  
QY 1766 ATTTGCTCGAGTCTCTGTGTACTAACAATGCTCAATATTTGTAAACTTGAATGGA 1825  
Db 1792 ACTGATTCGTGTATCTGTGTGTGACAAATGTGTGCTTATCTTCAATGTGATGATGGA 1851  
QY 1826 TCATCATCAACAACAGCAGTATTAAGAAAGCAATGTGTTTATGATGACCTTT 1885  
Db 1852 CCATTTACTTCAATGTGACAGCAAGCTCTTAGAAGCAATGTGCTTCAATGTGATCCAGC 1911  
QY 1886 ACTAGAAAGAGGTTGTGATGTACAGTCCCTCAAGANTTGAATGGAATGATCCGA 1945  
Db 1912 TCTAGAAAGGAAACTTGTATGTACAAATTTCCAAAGATGTATGATGATGATGATGAT 1971  
QY 1946 TGACCGATATGCTAACCGGAATGTGTCTTTTGTGATCAACATGAAAGGTTGATG 2005  
Db 1972 CGATCGATATGCTAATAGGAACATAGTCTTGTGATATCAACATGAAAGGTCTAGATG 2031  
QY 2006 TATTCAGGTCCAATTAATGTGTGTGATGATGTGATTTAGAGGAGCATTAATG 2065  
Db 2032 CATTCAAGGTCAGCTATATGTGGAACAGATGCTTTCAATGAGCAGGCTTGTATGG 2091  
QY 2066 TTATGATGCCCCCAAAACAAGAACCAATCAAGACTTGCAACTGTGCCCCAAGT 2125  
Db 2092 ATATGAT-----CTGTGTTTGAAGTGAAGTGTGAACTTAACAT 2134  
QY 2126 GTGCTTTGTGCTGTGCTTGTGCAATAGGAACAAAGAAAGCTACCAACCCAAAC 2185  
Db 2135 GTGTGTAAG---AGCTGTGTGTGAGAGAGAGAGAAAGAAAGTATATGATG 2190  
QY 2186 AGAGAGAAAAATTAATTTTCAAGAAAGAGAAACCAATCCCTGATATGCT 2245  
Db 2191 TCAAGCCGTA-----TTATGAGAGAAACGAATCTTCACTCCATCTTAACAT 2241  
QY 2246 TGTGAAATGACGAAGCTGTCCAGAGAGCTGAGAAATGAAAGCCGATATGTAATCA 2305  
Db 2242 GGAAGACATCAGAGAGGATTTGAAGGTATGAGAGATGAAAGTCACTGCTTAATGCCA 2301  
QY 2306 ACAAATTTAGAAAGAAATTTGSCAATCTTGTGTTTGTATCATCACTTCTGCA 2365  
Db 2302 GAGAAATTTAGAAAGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2361  
QY 2366 GAATGTGGAACCTTGAAGAGTGAAGTCTGTCTTCTTTTGAAGAGCTATACATGT 2425  
Db 2362 TCAAGGTGGCATACACCTTCAACAAACCGAGCTTCTTACTGAAGAGCTATCAATGT 2421  
QY 2426 CATTAGTTGTGTATGAGACAAAGACAGCTGGGGAAGAAAGATGAGCTGATCTATGG 2485  
Db 2442 TATCAGCTGTGGTACAGAGCAAAACTGATGAGGAAAGAGATGAGCTGATCTATGG 2481  
QY 2486 ATCAGTTACAGAAATATTTCAATGTTTCAAGATGATGATGATGATGATGATGATGAT 2545  
Db 2482 TTCAATTAACAGAAATTTCTGACTGTGTTTAAATGATGCAAGAGCTGCAATCAAT 2541  
QY 2546 TTACTGATACCTAAGCGGTTGATCAAAAGTTCTGCACTTGAATCTTTCAGATCG 2605  
Db 2542 CTACTGATGCAACAGACGACCTTGTTCAGAGGTTCTGCAACCAATCAATCTTTCAGATCG 2601  
QY 2606 TCTTCAACAGGTGCTTGGTGGGCTCTTGGGCTTATGAGATCTTTCAGCAATCATGG 2665  
Db 2602 TCTTAATCAGGTGCTCGTGGGCTCTTGGGCTCAGTGAATTCGTGCTTACACAACTTG 2661  
QY 2666 CCTCTTGTGTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2725  
Db 2662 TCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2721  
QY 2726 CTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785  
Db 2722 CACCATGTTTATCAATCAATCAATCTGTTCCGCTTATGCGCATATGTGTGCTTCTGCTAT 2781

OY	2786	CTGTTTATGACAGGGAAATTTATCACTCCAGAGCTGAATATATGTGGCACCCTGTGTT	2845
Db	2782	CTGCTCTTCAACATTAATTTATCATCTCGAATTTAGTAATATATGCTGGAAATGTTCTT	2841
OY	2846	CATGTCACTTTTATATCTGCATTTTTTGTACAGACATCCTAGAAAATGAGATGAGTGTGT	2905
Db	2842	CATTCTTCTTTTGTGCTCCATTTTGGCAACCTGTGTATATTGGAGCTCAGATGAGTGTGT	2901
OY	2906	TGGAATTGATACATGTGTGGAGAAATGACAGCTTCTGGGTCAATTGGAGTGTGTCTCACA	2965
Db	2902	TGGCATTTGMAATTTGTGTGAGAAAATGACAGATTTTGGGTATTGTGTGCACCTCTGCCCA	2961
OY	2966	CCCTCTTGTCTGTGTTCCAGGGACCTCTCAAGGTCTATACGTGTGTGATACAAGCTTTCAC	3025
Db	2962	TCTCTTTCCGGTGTTCCAGGGCTCTGCTAAGTGTGTGCTGGAGATTGATACCAATTTCAC	3021
OY	3026	CGTGACATCAAAAG--GTGAGATGATGAGAGATTTCCAGAGCTATATACATTCAAATG	3082
Db	3022	AGTTAACCCTAAAGGATCTGATGAGGATGGGGACTTGTCTAGAGCTATATGTTCAAAGTG	3081
OY	3083	GACTACCTTATTTGATACCTCCTACCACTTGTTCTTATTAACCTTCATTTGTTGTGTGC	3142
Db	3082	GACCGATTGTGCTACATCCCTCCGACCACTGTTCTTGTCAATTAACCTGTGCGAAATGTGTGC	3141
OY	3143	TGGCGTTTCAATGTGATCAATTAACGATATGAGTCAAGGGGCCCCCTTTGGGAAGCT	3202
Db	3142	AGGAATTCGTATGTGCATTTAACAGGGGCTACCAATCTCGGGTCCGCTCTTTGGAAAAGCT	3201
OY	3203	ATTCCTTGTCATTTGGGTGATTTGTCATCTTATTCCTTTCTTCCAAAGTTTGGTTGGAAG	3262
Db	3202	GTTCTTCTGCATCTGGGTGATCTCTTCATCTCAACCCCTTCTCAAGAGGCTTCATGGGAG	3261
OY	3263	GCAAAACAGACACCAACGATTTGTCATGCTGTGTCATTCCTGTGGCTTCAATCTTCTC	3322
Db	3262	GCAGAAACGCAAGCCAAACATCTGATCGTTTGGTCCATCTCCTTTGGCTATCTTCTC	3321
OY	3323	GCTCCTTTGGGTTGGGATTTGATCCCTTCTCT	3352
Db	3322	CTTGCTGTGGGTAAGATCGATCCCTTCTAT	3351

RESULT	15	
AAZ99506		
ID	AAZ99506	standard; DNA; 3725 BP.
XX		
XX	AAZ99506;	
AC		
XX		
DT	03-JUL-2000	(first entry)
XX		
DE		DNA encoding a maize cellulose synthase.
XX		
KW	Maize; cellulose synthase; stalk quality; strand; silage; cellulose;	
KW	transgenic plant; plant breeding marker; ss.	
XX		
OS		Zea mays.
XX		
FH	Key	Location/Qualifiers
FT	CDS	179..3397
FT		/*tag= a
FT		/product= "cellulose synthase"
FT		/note= "no termination codon given"
XX		
PN	WO200009706-A2.	
XX		
PD	24-FEB-2000.	
XX		
PF	16-AUG-1999;	99WO-US18760.
XX		
PR	17-AUG-1998;	98US-0096822.
XX		
PA	(PION-)	PIONEER HI-BRED INT INC.
XX		
PI	Dhugsa KS, Helentjaris TG, Bowen BA, Wang X;	

XX	MPi: 2000-224343/19.
XX	P-PSDB: AAY64112.
XX	New genes which encode maize cellulose synthase polypeptides in plants
XX	useful for modulating the expression of cellulose synthase in plants
XX	and to produce transgenic plants expressing the novel protein
XX	Claim 1; Page 121-126; 119pp; English.
XX	The present sequence encodes a maize cellulose synthase polypeptide.
XX	The cellulose synthase can be used for the improvement of stalk quality
XX	for improved stand or silage. It also provides an increased concentration
XX	of cellulose in the pericarp, hardening the kernel and improving its
XX	handling ability. The sequences are used to produce transgenic plants
XX	and seeds expressing the cellulose synthase. The polynucleotide is
XX	used for modulating, preferably increasing, the level of the synthase
XX	in a plant cell. The plants are preferably monocots. The polynucleotide
XX	is also used as a probe or primer in the detection quantitation or
XX	isolation of gene transcripts. The probes are useful in detecting
XX	deficiencies in the level of mRNA in screenings for desired transgenic
XX	plants, for detecting mutations in the gene, for monitoring upregulation
XX	of expression or changes in enzyme activity in screening assays of
XX	compounds, for detection of any number of allelic variants of the gene,
XX	or for use as molecular markers in plant breeding programs. The
XX	isolated nucleic acids of the present invention can also be used for
XX	recombinant expression of their encoded polypeptides or for use as
XX	immunogens in the preparation and/or screening of antibodies. The
XX	proteins can be employed in assays for enzyme agonists or antagonists
XX	of enzyme function or for use of immunogens or antigens to obtain
XX	antibodies specifically immunoreactive with a protein.
XX	Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;
XX	Query Match 35.0%; Score 1378.4; DB 21; Length 3725;
XX	Best Local Similarity 66.4%; Pred. No. 0;
XX	Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9
QY	119 CGGCGGAAGTGAAGGAGGAAGGAGCATGAGAGCGCGCGCTGTGGCCGCTCCCA 178
DB	154 CGGTGTGTGTGAGAGGAGCGCGGAGATGGCGGCAACAAGGGGATGTGGCAGGCTTCA 213
QY	179 CAACCGCAACGAGCTCTGTCTATTCGCGCGGACGGCA-----TCCGCGGCGGAAGCC 232
DB	214 CAACCGCAACGAGTTCGTCTATTCGCGCGGACGGCAACGCGCTGTCCGCGCTAAGCC 273
QY	223 GCGCGGAGAGCAAGACGGGACAGGTGTGCAGATTGTGCGGCGACGAGTCCGCTTGCCCC 292
DB	274 CACGGAAGGTGCGAATGGGACAGGTGTGCAGATTGTGCGGACGAGTGTGGCTTTGACG 333
QY	233 CGGCGGAGACCCCTTCTGTGGCGGTGCAAGAGTGCCTTCCCGCTGTGCGGGAATGTCTA 352
DB	334 CACGTGTATGTCTTTTGTGCTGCAATGAGTGTGCTTTCCCTGTCTGCGCGCTTGCTCA 393
QY	353 CGAATACAGAGCGCGGAGGAGGACGACGAATCTGCCCCCAATGACATCTGATACAGCG 412
DB	394 TGAGTACAGAGCGCAAGGAAGGAACCAATGTCTGCCCTCAATGACCAAGACTAACAAGAG 453
QY	413 CCTCAAGGCGTCCCAAGCTGTGACCGGTGACGAGAGAGAGACGCGGCTGTGATGACTGGA 472
DB	454 ACAGAAAGTGTAGCCCTCGAATTCATGTGTGATGATAGGAGGAAGATGTTGATGACTTGA 513
QY	473 CAACGAGTTCAACTGGGACGGGACATGACTGCGATCTGTGGCGGAGTCCATCTCTACGG 532
DB	514 CAATGAATTTCAACTTAA-----GCAAGGCAATGGGAAGG 549
QY	553 CCACATGAGCTACGGCGGTGAGAGTGAACCTTAATGGCGCGCCACAAGCTTTCAGCTCA 592
DB	550 CCGAAGTGTGAGAGCTTCAAGAGAGATGACGTGATCTGTCTTATCTGCTCGGACATGACCC 609
QY	593 CCGCAAT---GTTCATCTCTCAACCAAGGGAATGTGTGATGATCAATCCACCGGAGCA 649
DB	610 ACACCATGAGATTCACGCGCTTCAACAGTGAACAAAGATATCTGGAGAGATCTCTGATGC 669

```
Qy 650 GCAGCGCTGTCCTTTTCATGCGTGGGGGAAAGATATCCCTTCTTA 709
Db 670 ATCCCTGAGCCGTATCTATCCGAGT-----CCAAATCGAGCTA 711
Qy 710 TGGCGATCCAGCTTACCTGTGCAACCGAGTCTATGGACCCATCGAAGATCTTGTC 769
Db 712 TGTGATTCGAAGCTTCCAGTTCCTGTAGAGATGTGGAACCCCTCGAAGGACTTGAATTC 771
Qy 770 ATATGGTATGTAGTGTGCTTGGAAAGAAAGATGGAAATTTGAAGCAGACAGA 829
Db 772 CTATGGGCTTAATAGTGTGACTGGAAAGAAAGATTGAGAGCTGGAGGGTTAAACGA 831
Qy 830 GAGATGACCAAGACGGGAAATGATGCTGTGTGATATGCTG----- 873
Db 832 CAAAAATATGTGGCAAGTACTAATTAATTCAGAGGCTAGAGAGACATGAGGGGAC 891
Qy 874 -----ACGATGCTGATCTACCACTAATGGATGAAGCAAGCAACATGTCAGGAA 925
Db 892 TGGCTCAATGGAAGATATGCAAAATGTGATGATGACGCCCTACCTTTGAGCGCAT 951
Qy 926 AATTCACTTCATCAAGCCAGATTAATCAATATAGATGATTAATCATTAATTCGCTGT 985
Db 952 TGTGCAATTTCTTCAACAGCTCAACCTTTACCGATATGATTCATTCCTGCTTAT 1011
Qy 986 GGTTTGGGGTTCTTCTTCACTACCGAGTATGATCCGTTGAATGATGCAATTTGCTT 1045
Db 1012 CATCTGTGCTTCTTCAATATCGTATCGATCATCAGTGGGTATGCTATATGAT 1071
Qy 1046 GTGCTCATATCTGTTATCTGTAATCTGTTGGTTCCATGCTTGGATCTTGTACAT 1105
Db 1072 GTGCTGATATCTGTTATCTGTAGAGTGTGTTGCTTGTGCTTGTGAGTCAAGTT 1131
Qy 1106 CCCAAAGGTGTCTCTATTTAGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGTTGCA 1165
Db 1132 CCGAAATGTATTCATCAACCGTGAACATATCTGACAGGCTGTGATGAGTATGA 1191
Qy 1166 CAAGAAAGCCAGCCATCTCACTTGTCCAAATGATTTCTTTGTCACTACGTTGATCC 1225
Db 1192 TAGAGAGGAGGACATCACAGCTGGCTCCCATTTGATGTCTTTGTCACTACGTTGCC 1251
Qy 1226 CTTAAGGAACCTCTTGTGTCACAAATACTGTTATCTATCTTCCCTTGGGTGATTA 1285
Db 1252 ATTAAGGAACCTTCCATGATCACAGCCAACTGTTTGTTCATCTTGTGTGATTA 1311
Qy 1286 TCCGTGTATAGGTTTCTTCTATGTTTCTGATGATGCTGCAATGCTAAAGTTGA 1345
Db 1312 CCGTGTGACAAAGTGTATGCTATGTTTCTGATGATGCTCAGCTATGCTGACTTTGA 1371
Qy 1346 AGCATTTATCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTTTGTGCAACGGTACA 1405
Db 1372 GTCTCTCTGAAACCTGCGAATTTGCTAAGAAAGTGGGTTCCCTTTGTGAAGACAA 1431
Qy 1406 TATTGAACCTGCGCTCGAGGTGTACTTCCACAGAAATATGACTTCTGAAAGACA 1465
Db 1432 TATTGAACCAAGAGCTCGAATTTTACTTTGCTCAAAAATAATGATTAACCTGAAAGACA 1491
Qy 1466 GGTGGCAGCAACTTTGTTAGGAGAGAGACATGAAAGAGATGAGATTCAA 1525
Db 1492 AATTCAACTTCAATTTGTTAAGAAAGACGCAATGAAAGAGATGAGAAATTCAA 1551
Qy 1526 GGTGAGATCAATGCTTAGTTGCCAAAGCCCAAGAAATTCCTGAAGAGATGACAT 1585
Db 1552 AATAAGATCAATGCTTGTGTTGCAAGACAGAAAGTGTCTGAAGAGGGGTGACAT 1611
Qy 1586 GCAAGATGGAACCCCTGTGCTGGAAGAAATGTTCTGTATCATCTGAAATGATTCAGT 1645
Db 1612 GGTGATGGAAGTGTGTGGCTGTGGGAATTAACCTAGGACATCTGTGCAATTCAGT 1671
Qy 1646 CTTCCTTGGCCAAAGCGAGGCTTGACTGTGAGGGAATGAACTGCCAGATTTGTTA 1705
Db 1672 GTTCTTGGGGCAAGTGTGGGCTTTGACATGATGGAATTAATTAACGACTTGTCTA 1731
Qy 1706 TGTTTTAGAGAGAAAGACAGGCTATTAACATCAATAAAGAAAGCTGTGCTATGATGC 1765
Db 1732 TGTCTCTGTAAGAAAGACAGGCTTTTCAAGCATCAAAAGAGCTGTGCTAATGAATGC 1791
Qy 1766 ATGTGTCGAGTCTGTGCTGTACTTAACAAATGTCTCATATTTGTAACTTGTATGTA 1825
Db 1792 ACTGATTCGTGTATCTGTGCTGTGACAAATGTGTCTATCTCTCAATGTGATGTGTA 1851
Qy 1826 TCACTACATCAACACAGCAAGGCTATTAAGAAAGCAATGTGTTTATGATGACCTTT 1885
Db 1852 CCAATTAATCAATAGCAGCAAGCTCTTAAGAAAGCAATGTGCTCATGATGATCCAGC 1911
Qy 1886 ACTAGAAAGAAAGTTTGTCTATGATCAAGTTCCTCAAAAGTTTGTATGATGATGCGCA 1945
Db 1912 TTAAGAAAGAAACCTTGTATGATCAATTTTCAAAAGATTTGATGAGCACTTGCA 1971
Qy 1946 TGAACGATATGCTTAACCGGAATGTGCTTTTGTATATCAACATGAAGGTTTGTATG 2005
Db 1972 CGATCGATATGCTTAATGGAACATAGTCTTCTTGTATATCAATGAAGGCTTATAGTG 2031
Qy 2006 TATTCAAGGTTCAATTTATGTTGTATCTGATGTGATATTTAGAAGGACGATTAATAG 2065
Db 2032 CATTCAGGTTCCAGTCTATGTGGAAACAGATGCTGTTCAATAGGACGCTTTGTATGG 2091
Qy 2066 TTATGATGCCCCCAAAAGCAAGCCATCAAGGACTTGAACCTGCTGGCCAAAGT 2125
Db 2092 ATATGAT-----CCTGTTTATCTGAAGCTGATCGGAACCTTAATCT 2134
Qy 2126 GTGCTTGTGCTGTGCTTGTGCAATAGAAAGCAAAAGAAAGCTACCAACCCCAAC 2185
Db 2135 GTTGTTAAG---AGCTGTGTGTGATAGAAAGAAAGAAAGCAAGTATATGATAG 2190
Qy 2186 AGAAGAAAGAAAGTATTAATTTTTCAGAAAGAAAGAACCAATCCCTGCTATATGCT 2245
Db 2191 TCAAGCCGTA-----TTATGAAGAAAGAAAGATCTTCAAGTCCCATCTTAAAT 2241
Qy 2246 TGTGAAATGAGAGAGCTGCTCAAGAGCTGAGATGAAGAAAGCCGCTATTTGAATCA 2305
Db 2242 GGAAGACATGAGAGAGGATTAAGAGTTATGAGATGAAGAAAGGCTATATGTTGCCA 2301
Qy 2306 ACAAATTTGAAAGAAATTTGGCCAACTTCTGTTTGTATGATCCACACTTCGCA 2365
Db 2302 GAGAAATTTGAGAAAGCTTTGTGTGATGCTTCCAACTTCAATTTGATGATCCACTTATGAC 2361
Qy 2366 GAATGTGGAACCTTGAAGAGTGCAGAGTCTGCTTCTTTTGAAGAACTATACATGT 2425
Db 2362 TCAAGGTGGAATCCACTTCAACAAACCAAGCTTCTCTGTAAGAACTATACATGT 2421
Qy 2426 CATTAATGTTGTATGATGAACAAGACAGACTGGGGAAGAAAGATGCTGATCTATGG 2485
Db 2422 TATCAGCTGTGGGTACAGAGCAAACTGAATGGGGAAGAAAGATTTGGCTGATCTATGG 2481
Qy 2486 ATCAGTTACAGAGATTTCTAATGTTTCAAGATGCAATGTCATAGTGTGGCGTAAT 2545
Db 2482 TTCAATGACAGAGATTTCTGACTGGGTTTAAATGATGCAAGAGGCTGGCAATCAAT 2541
Qy 2546 TTAATGATACCTTAACCGGTTGATCAAGAGTTCTGCACTCTGATCTTTTCAATG 2605
Db 2542 CTACTGATGCAACAGACACTTGTTTCAAGGGTTCTGCAACATCAATCTTTCTGATCG 2601
Qy 2606 TCTTCAACAGGCTTGTGGGTCTCTGGGCTCTATTTAGATCTTTCAAGCAATCAATG 2665
Db 2602 TCTTAATCAGGTGCTCGTGGGCTCTGGGCTCAATGGAATTTCTGTAGCAGACATGG 2661
Qy 2666 CCGCTTGTGATGGGATGATGGGCGGTGAATTTTGTGAAGATTTTCTAATCA 2725
Db 2662 TCTTATATGATGATGCTCAATGAGCGATTTGAAGCTTTGAGAGAGCTGTGCTTACATTA 2721
Qy 2726 CTCCATGATGATCTTGAACATTAATCCCTCTTGGCTTACATGATCAATTTGCTGCAAT 2785
Db 2722 CACCATGTTTATTCATCAATCAATCTGTTCCGCTATATGCTCTATGCTGTCTGCTAT 2781
Qy 2786 CTGTTTATGACAGGGAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGCTT 2845
```

```
Db 2782 CTGCTCTTACCAATAAATTTATCTCTGAGATTAGTAATATATGCTGGAATGTTCTT 2841
Qy 2846 CATGTCACTTTTATCTGATGATTTTGTCTACGAGCATCTAGAAATGAGATGAGTGTGT 2905
Db 2842 CATTCCTTTTCTCTCCATTTTCCCACTGGTATATTGAGCTCAGATGAGATGAGTGTGT 2901
Qy 2906 TGGAAATTGATGACTGTGTGAGAGAAATAGAGATTTCTGGGTCAATTGAGAGTGTGTCTCACA 2965
Db 2902 TGGCAATTGAAGATGTGTGAGAGAAATAGAGATTTTGGGTATTTGTGGCACTGTGCCCA 2961
Qy 2966 CCTCTTGTCTGTGTCCAGGAACTTCTCAAGGTCATAGCTGTGTGTGATACAAAGCTTCAC 3025
Db 2962 TCTCTTCGCGGTCTTCCAGGTCGTGCTGAAGTGTGTGCTGGATTTGATACCACTTCAC 3021
Qy 3026 CGTGACATCAAAAG--GTGAGATGATGAGAGATTTCTCAGAGCTATATACATTCAATG 3082
Db 3022 AGTTACCTCAAAAGGCATCTGATGAGAGATGAGCACTTGTCTGAGCTATATGTGTCAAGTG 3081
Qy 3083 GACTACCTTATTTGATACCTCTTACCACTTGTCTTATTTGAATTCATTGTGTGTGTGCGC 3142
Db 3082 GACCAAGTTTGCATCTCCCTCCGACCACTGTTCTTGCATTAACTGTGTGGAATGTGTGC 3141
Qy 3143 TGGCGTTTCAAAATGCGATCAATPACGATATGAGTCATGAGGCCCCCTCTTTGGGAAGCT 3202
Db 3142 AGGAATTTTGTATGCTATTACAGCGGCTACCAATCTGTGGGTCCGCTCTTTGGAAAGCT 3201
Qy 3203 ATTCTTTGCAATTTGGGTGATTTGTCATCTTTATCCCTTTCTCAAAAGTTTGGTTGGAAG 3262
Db 3202 GTTCTTCTCGATCTGGGTGATCTCTCATCTCTACCCCTTCTCAAGGTCATGAGGCGAG 3261
Qy 3263 GCMAAACAGAGACCAACGATTTGTCATGCTGTGTGTCATTCTGTGCTTCATCTTCTC 3322
Db 3262 GCMAAACCGACCGCAACAATCGTCATCGTTTGTGTCTATCTTGGCGTATCTTCTC 3321
Qy 3323 GCTCCTTTGGGTTCGAGTTGATCCCTTCTCT 3352
Db 3322 CTGCTGTGGGTGAAGATCGATCTTTCAT 3351
```

Search completed: December 15, 2003, 18:48:39  
Job time : 979 secs



Dh 121 GCGAAGTGAAGGGAGGAAAGCAATGAGGCGAGCGCGGCTGTGGCCGGCTCCACA 180  
Qy 181 ACCGCAACGAGCTGTGTGATCGCGCGAGCCGCGATCCCGGCGAGACCGCGGG 240  
Dh 181 ACCGCAACGAGCTGTGTGATCGCGCGAGCCGCGATCCCGGCGAGACCGCGGG 240  
Qy 241 AGCAGAAACGGGCAAGTGTGCAATTTGGCGCGACGATCGCTTGGCCCCGGCGGG 300  
Dh 241 AGCAGAAACGGGCAAGTGTGCAATTTGGCGCGAGCGATCGCTTGGCCCCGGCGGG 300  
Qy 301 ACCGCTTGTGGCGTGAAGAGAGTGTGCTTCCCGCTGCGCGGAGCTGTGAGAAATACG 360  
Dh 301 ACCGCTTGTGGCGTGAAGAGAGTGTGCTTCCCGCTGCGCGGAGCTGTGAGAAATACG 360  
Qy 361 AGCGCGGAGAGGCGCAGAACTGCCAGTGAAGAGTGTGAGTACAGAGCGCTCAAGG 420  
Dh 361 AGCGCGGAGAGGCGCAGAACTGCCAGTGAAGAGTGTGAGTACAGAGCGCTCAAGG 420  
Qy 421 GCTGCCAACGTGTGACCGGTGAAGAGAGAGAGCGCGCTGTGATGACCTTGGACAAGT 480  
Dh 421 GCTGCCAACGTGTGACCGGTGAAGAGAGAGAGCGCGCTGTGATGACCTTGGACAAGT 480  
Qy 481 TCAACTGGAGCGGCGCACTGCGAGTCTGTGGCGGAGTCAATGCTTAAGCGCAATGA 540  
Dh 481 TCAACTGGAGCGGCGCACTGCGAGTCTGTGGCGGAGTCAATGCTTAAGCGCAATGA 540  
Qy 541 GCTACGCGCGGTGAGGTGACCTTAATGAGCGCGCAAGGCTTTTCAGCTCAACCCCATG 600  
Dh 541 GCTACGCGCGGTGAGGTGACCTTAATGAGCGCGCAAGGCTTTTCAGCTCAACCCCATG 600  
Qy 601 TTGCACTCTTCACCAACGGGCAATGTGTGATGACATCCACCGGAGACAGCGCGTGG 660  
Dh 601 TTGCACTCTTCACCAACGGGCAATGTGTGATGACATCCACCGGAGACAGCGCGTGG 660  
Qy 661 TGCCTTCTTTCATGGGAGTGTGGGGGAAAGAGATACATCCCTTCTTATGGGATCCCA 720  
Dh 661 TGCCTTCTTTCATGGGAGTGTGGGGGAAAGAGATACATCCCTTCTTATGGGATCCCA 720  
Qy 721 GCTTACCTGTGCAACCCAGGTCTATGAGCCCATCCAGAGATCTTGTGTCATATGGGTATG 780  
Dh 721 GCTTACCTGTGCAACCCAGGTCTATGAGCCCATCCAGAGATCTTGTGTCATATGGGTATG 780  
Qy 781 GTAGTGTGCTTGGAGAGAAACGATGAGAAATTTGGAAGCAAGACAGAGAGATGAC 840  
Dh 781 GTAGTGTGCTTGGAGAGAAACGATGAGAAATTTGGAAGCAAGACAGAGAGATGAC 840  
Qy 841 AGACGGGAAATGATGTGTGTGATGATGTGTGATGATGTGTGATGATGTGTGATG 900  
Dh 841 AGACGGGAAATGATGTGTGTGATGATGTGTGATGATGTGTGATGATGTGTGATG 900  
Qy 901 AAGCAGACAACTGTCCAGGAAATTTCCACTTCATCAAGCCAGATTAATCCATATA 960  
Dh 901 AAGCAGACAACTGTCCAGGAAATTTCCACTTCATCAAGCCAGATTAATCCATATA 960  
Qy 961 GAGATGATTAATTAATTCGGCTTGTGGTGTGGGATCTTCTTCACTACCGAGTATGC 1020  
Dh 961 GAGATGATTAATTAATTCGGCTTGTGGTGTGGGATCTTCTTCACTACCGAGTATGC 1020  
Qy 1021 ATCCGAGTGAATGATGATTTGCTTGTGGCTCATATGCTTATGTGAAATCTGTTTG 1080  
Dh 1021 ATCCGAGTGAATGATGATTTGCTTGTGGCTCATATGCTTATGTGAAATCTGTTTG 1080  
Qy 1081 CCAATGCTTGGATCTTGTATCAATTTCCAAAGTGTGCTTCCATTTGAGAGAGAGCTTACC 1140  
Dh 1081 CCAATGCTTGGATCTTGTATCAATTTCCAAAGTGTGCTTCCATTTGAGAGAGAGCTTACC 1140  
Qy 1141 TAGACCGGCTGTCACTGAGGTTCGACAAAGAGGCGACGATCTCAACTTGTCTCAATTG 1200  
Dh 1141 TAGACCGGCTGTCACTGAGGTTCGACAAAGAGGCGACGATCTCAACTTGTCTCAATTG 1200  
Qy 1201 ATTTCTTTGAGAGAGGTGATCCCTTAAAGAACTCTTGGTGTACAACTTCTG 1260  
Dh 1201 ATTTCTTTGAGAGAGGTGATCCCTTAAAGAACTCTTGGTGTACAACTTCTG 1260

Qy 1261 TTCTATCTATCCCTTGGGTGATATTCCTGTGATTAAGGTTTCTTGTCTATGTTCATG 1320  
Dh 1261 TTCTATCTATCCCTTGGGTGATATTCCTGTGATTAAGGTTTCTTGTCTATGTTCATG 1320  
Qy 1321 ATGTGCTGCAATGCTTAAGCTTTGAAGCTTATCTGAACATCTGAATTTGCAAGAAAT 1380  
Dh 1321 ATGTGCTGCAATGCTTAAGCTTTGAAGCTTATCTGAACATCTGAATTTGCAAGAAAT 1380  
Qy 1381 GGGTTCCTTTCGCAACCGGTACAAATATTGAACCTCGCGCTCAGAGTGTACTTCAAC 1440  
Dh 1381 GGGTTCCTTTCGCAACCGGTACAAATATTGAACCTCGCGCTCAGAGTGTACTTCAAC 1440  
Qy 1441 AGAAGATAGACTTCTTGAAGCAAGGTGCGCAACTTTGTAGGAGAGAGAGCA 1500  
Dh 1441 AGAAGATAGACTTCTTGAAGCAAGGTGCGCAACTTTGTAGGAGAGAGAGCA 1500  
Qy 1501 TGAAGAGAGATATGAGAAATCAAGGTGAGAAATGCTTATGTTGCAAGGCGACA 1560  
Dh 1501 TGAAGAGAGATATGAGAAATCAAGGTGAGAAATGCTTATGTTGCAAGGCGACA 1560  
Qy 1561 AAGTTCCTGAAGAGATGAGCAATGACAGATGAAACCCCTGCGCTGGAACAAATGTT 1620  
Dh 1561 AAGTTCCTGAAGAGATGAGCAATGACAGATGAAACCCCTGCGCTGGAACAAATGTT 1620  
Qy 1621 GTGATATCTTGAATGATTCAGGTCTTCTTGGCCAAAGCGAGGCTTGACTGTGAG 1680  
Dh 1621 GTGATATCTTGAATGATTCAGGTCTTCTTGGCCAAAGCGAGGCTTGACTGTGAG 1680  
Qy 1681 GAAATGAACTGCGAGATTGTTTATGTTTCTTGAAGAGAAAGACAGGCTTAACATC 1740  
Dh 1681 GAAATGAACTGCGAGATTGTTTATGTTTCTTGAAGAGAAAGACAGGCTTAACATC 1740  
Qy 1741 ATAGAAGACGTGTGCTATGAATGCAATGTCGAGTCTTCTGTACTTAACAAATGCTC 1800  
Dh 1741 ATAGAAGACGTGTGCTATGAATGCAATGTCGAGTCTTCTGTACTTAACAAATGCTC 1800  
Qy 1801 CATATTTGTTAACTTGAATGTGATCACTACATCAACAGCAAGGCTATTAAGGAAG 1860  
Dh 1801 CATATTTGTTAACTTGAATGTGATCACTACATCAACAGCAAGGCTATTAAGGAAG 1860  
Qy 1861 CAATGTGTTTATGAGGAGCCCTTATCTAGAAAGAGGTTGCTATGACGTTCCCTC 1920  
Dh 1861 CAATGTGTTTATGAGGAGCCCTTATCTAGAAAGAGGTTGCTATGACGTTCCCTC 1920  
Qy 1921 AAAGATTTGATGGATGATTCGCATGACGATATGCTAACCGGAATGTGTCTTTTG 1980  
Dh 1921 AAAGATTTGATGGATGATTCGCATGACGATATGCTAACCGGAATGTGTCTTTTG 1980  
Qy 1981 ATATCAACATGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Dh 1981 ATATCAACATGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Qy 2041 TATTTGAAGGAGGAGGATTAATGTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Dh 2041 TATTTGAAGGAGGAGGATTAATGTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Qy 2101 GAGCTTGAACCTGCTGCGCAAGTGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTG 2160  
Dh 2101 GAGCTTGAACCTGCTGCGCAAGTGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTG 2160  
Qy 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAGTTATTTTCAAGAAAGAG 2220  
Dh 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAGTTATTTTCAAGAAAGAG 2220  
Qy 2221 AGAAGCAATCCCTGATATGCTTGTGGAATTTGAAGAGCTGCTCAAGAGCTGAGA 2280  
Dh 2221 AGAAGCAATCCCTGATATGCTTGTGGAATTTGAAGAGCTGCTCAAGAGCTGAGA 2280  
Qy 2281 ATGAAGAGCGGATATTTGAATCAACAAATTTAGAAAGAAATTTGGCAATCTTCTG 2340  
Dh 2281 ATGAAGAGCGGATATTTGAATCAACAAATTTAGAAAGAAATTTGGCAATCTTCTG 2340





121 GCGAAGTGGAGGGAGGAAAGCATGAGGCGAGCCGGGCTGGTGGCCGGCTCCACA 180  
121 GCGGAAGTGGAGGGAGGAAAGCATGAGGCGAGCCGGGCTGGTGGCCGGCTCCACA 180  
181 ACCGCAAGAGCTCGTGCATCGGCGCGAGCGGATCCGGGGCGGAAGCCGGCGGG 240  
181 ACCGCAAGAGCTCGTGCATCGGCGCGAGCGGATCCGGGGCGGAAGCCGGCGGG 240  
241 AGCAGAACGGGAGGATGTGCCAGATTTGGCGGAGAGAGCGTGGGCTGGCCCGGGGG 300  
241 AGCAGAACGGGAGGATGTGCCAGATTTGGCGGAGAGAGCGTGGGCTGGCCCGGGGG 300  
301 ACCCTTGGTGGCGTGCACAGAGTGGCTTCCCGTCTGGCCGGAGCTGTACGAATAAG 360  
301 ACCCTTGGTGGCGTGCACAGAGTGGCTTCCCGTCTGGCCGGAGCTGTACGAATAAG 360  
361 AGCGCGGGAGGGAGCGAGAACTGCCCCAGTGGAGACTGATACAAAGCCGCTCAAG 420  
361 AGCGCGGGAGGGAGCGAGAACTGCCCCAGTGGAGACTGATACAAAGCCGCTCAAG 420  
421 GCTGCGAACGTGTGACCGGTGACGAGAGAGAGAGCGGCGTGCAGTACCTGGAACAAGT 480  
421 GCTGCGAACGTGTGACCGGTGACGAGAGAGAGAGCGGCGTGCAGTACCTGGAACAAGT 480  
481 TCACCTGGAGCGGCGCATGACTGCGAGTCTGTGGCGGAGTCATGCTTACGCGCACATGA 540  
481 TCACCTGGAGCGGCGCATGACTGCGAGTCTGTGGCGGAGTCATGCTTACGCGCACATGA 540  
541 GGTAGGGCGGTGGAGGTGACCTTAATGGGGCGGCAAGACTTCCAGCTCAACCCCAATG 600  
541 GGTAGGGCGGTGGAGGTGACCTTAATGGGGCGGCAAGACTTCCAGCTCAACCCCAATG 600  
601 TTCCACTCCTCAACCAAGGCAAAATGTGTGATGATCCCAACGAGAGAGAGCGGCTGG 660  
601 TTCCACTCCTCAACCAAGGCAAAATGTGTGATGATCCCAACGAGAGAGAGCGGCTGG 660  
661 TGCCTTCTTTCATGGGTGGTGGGGAAGAGATACATCCCTTCTTAATGCGAATCCCA 720  
661 TGCCTTCTTTCATGGGTGGTGGGGAAGAGATACATCCCTTCTTAATGCGAATCCCA 720  
721 GGTACCTGTGACCCGAGTCTATGAGACCCATCCAGATCTTCTGTGATATGGGATG 780  
721 GGTACCTGTGACCCGAGTCTATGAGACCCATCCAGATCTTCTGTGATATGGGATG 780  
781 GTAGTGTGCTTGGAGAGAACGATGAGAAATTTGGAAGCAGAGAGAGAGATGCAAC 840  
781 GTAGTGTGCTTGGAGAGAACGATGAGAAATTTGGAAGCAGAGAGAGAGATGCAAC 840  
841 AGACGGGGAGATGATGGTGGTGGTGTGATGTGACGATGCTGATACCACTAATGATG 900  
841 AGACGGGGAGATGATGGTGGTGGTGTGATGTGACGATGCTGATACCACTAATGATG 900  
901 AAGCAAGACAAACATGTCCAGAGAAATTTCCACTTCCATCAAGCCAGATTAATCCATTA 960  
901 AAGCAAGACAAACATGTCCAGAGAAATTTCCACTTCCATCAAGCCAGATTAATCCATTA 960  
961 GGATGATTATCATTAATTCGGCTGTGGTTTGGGGTCTTCTTCCACTACGAGATGATG 1020  
961 GGATGATTATCATTAATTCGGCTGTGGTTTGGGGTCTTCTTCCACTACGAGATGATG 1020  
1021 ATCCGGTGAATGATCATTTGCTTGTGGCTCATATCTGTATCTGTGAAATCTGATTTG 1080  
1021 ATCCGGTGAATGATCATTTGCTTGTGGCTCATATCTGTATCTGTGAAATCTGATTTG 1080  
1081 CCATGCTTGGATCTCTGATCAATTCGCAAGTGGTCCCTAATTAAGAGAGACTTACC 1140  
1081 CCATGCTTGGATCTCTGATCAATTCGCAAGTGGTCCCTAATTAAGAGAGACTTACC 1140  
1141 TAGACCGGCTGTCACTGAGGTTGACAGAGAGAGGCGAGCATCTCACTTGTCTCCAAATTG 1200  
1141 TAGACCGGCTGTCACTGAGGTTGACAGAGAGAGGCGAGCATCTCACTTGTCTCCAAATTG 1200

1201 ATTCTTGTGAGTGAAGGTTGATGCCCTTAAGAAAGAACTCTTGTGTGACAAACAATATCTG 1260  
1201 ATTCTTGTGAGTGAAGGTTGATGCCCTTAAGAAAGAACTCTTGTGTGACAAACAATATCTG 1260  
1261 TTCTATCATCTCTTGGGTGATATCTGTGTGATTAAGTTTCTTGTATATGTTCTGATG 1320  
1261 TTCTATCATCTCTTGGGTGATATCTGTGTGATTAAGTTTCTTGTATATGTTCTGATG 1320  
1321 ATGGTGTGCAATGCTAACGTTTGAAGCATTAATCTGAACATCTGAATTTTGCAAGAAAT 1380  
1321 ATGGTGTGCAATGCTAACGTTTGAAGCATTAATCTGAACATCTGAATTTTGCAAGAAAT 1380  
1381 GGGTTCCTTCTTGCAGAACGATTAATTTGAACCTCGCGCTCCAGAGTGGTATCTTCAAC 1440  
1381 GGGTTCCTTCTTGCAGAACGATTAATTTGAACCTCGCGCTCCAGAGTGGTATCTTCAAC 1440  
1441 AGAAGATGACTACTTGAAGAAAGCAAGGTGGAGCAAACTTTGTAAGGAGAGAGAGCA 1500  
1441 AGAAGATGACTACTTGAAGAAAGCAAGGTGGAGCAAACTTTGTAAGGAGAGAGAGCA 1500  
1501 TGAAGAGAGATGAGAAATTCAGGTGAGAAATCAATGCTTAATGTTGCCAAGCCGAGA 1560  
1501 TGAAGAGAGATGAGAAATTCAGGTGAGAAATCAATGCTTAATGTTGCCAAGCCGAGA 1560  
1561 AAGTTCCTGAAGAGATGAGCAATGCAAGATGGAACCCCTGGCTCGAGAAACAATGTT 1620  
1561 AAGTTCCTGAAGAGATGAGCAATGCAAGATGGAACCCCTGGCTCGAGAAACAATGTT 1620  
1621 GTGATCATCTCTGAGTATTCAGGTCTTCCCTGGCCAAAGCGAGGCTTGAATGATGAGG 1680  
1621 GTGATCATCTCTGAGTATTCAGGTCTTCCCTGGCCAAAGCGAGGCTTGAATGATGAGG 1680  
1681 GAAATGAACTGCGACGATGGTATATGTTCTAGAGAAACGACGAGGCTATACCATC 1740  
1681 GAAATGAACTGCGACGATGGTATATGTTCTAGAGAAACGACGAGGCTATACCATC 1740  
1741 ATAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
1741 ATAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
1801 CATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
1801 CATATTTGTTAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
1861 CAATGTGTTTATGATGACCTTTTACTAGAAAGAGTTTGTCTATGATCACTTCCCTC 1920  
1861 CAATGTGTTTATGATGACCTTTTACTAGAAAGAGTTTGTCTATGATCACTTCCCTC 1920  
1921 AAAATTTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980  
1921 AAAATTTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980  
1981 AATACCATGAAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
1981 AATACCATGAAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
2041 TATTAGAGGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
2041 TATTAGAGGCAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
2101 GGACTTGCACATGCTGGCCCAAGTGGCTTTTGTGCTGTGCTTGGCAATGAGAAC 2160  
2101 GGACTTGCACATGCTGGCCCAAGTGGCTTTTGTGCTGTGCTTGGCAATGAGAAC 2160  
2161 AAAAGAGACTACCAAAACCCAAACAGAGAGAAAGTTATTTTTCAGAGAAAG 2220  
2161 AAAAGAGACTACCAAAACCCAAACAGAGAGAAAGTTATTTTTCAGAGAAAG 2220  
2221 AGAACCAATCCCTGCAATATGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 2280  
2221 AGAACCAATCCCTGCAATATGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 2280  
2281 ATGAAAAGGCGGATTTGTAATCAACAAAATTTAGAAAAGAAATTTGGCAATCTTCTG 2340



Db 1 CTTTCCCTCGTGGGTGGGCGGGCTGGGCTGGGTGGTGAAGAAACAATCGGGGG 60  
Qy 61 ATGAGATCTGCTGCTAGAGTGAAGAGACTACGGTCAGTATCTCTGCTTGTGGCG 120  
Db 61 ATGAGATCTGCTGCTAGAGTGAAGAGACTACGGTCAGTATCTCTGCTTGTGGCG 120  
Qy 121 GGGGAAGTGAAGGGAAGAGAGTGAAGAGAGCGGGGCTGGTGGCGGCTCCCA 180  
Db 121 GGGGAAGTGAAGGGAAGAGAGTGAAGAGAGCGGGGCTGGTGGCGGCTCCCA 180  
Qy 181 ACCGCAACGAGCTCGTCTCATCCGCGGAGAGCGGATCCCGGCGGAAGCGCGCGG 240  
Db 181 ACCGCAACGAGCTCGTCTCATCCGCGGAGAGCGGATCCCGGCGGAAGCGCGCGG 240  
Qy 241 AGCAGAACGGGCAAGGTGTCCAGATTTGGCGGAGAGAGCTGGGCTTGTGGCGGG 300  
Db 241 AGCAGAACGGGCAAGGTGTCCAGATTTGGCGGAGAGAGCTGGGCTTGTGGCGGG 300  
Qy 301 ACCCGTGGTGGCGGAGAGAGTGGCGCTTCCCGCTGGCGGGAGCTGTAGAAATCG 360  
Db 301 ACCCGTGGTGGCGGAGAGAGTGGCGCTTCCCGCTGGCGGGAGCTGTAGAAATCG 360  
Qy 361 AGCGCGGAGGAGGAG 420  
Db 361 AGCGCGGAGGAGGAG 420  
Qy 421 GCTGCCAACGTGTGACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Db 421 GCTGCCAACGTGTGACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Qy 481 TGAACGTGGAGCGGCACTGAGCTGGAGTGGCGGAGTCAATGCTCTAGAGGCAATGA 540  
Db 481 TGAACGTGGAGCGGCACTGAGCTGGAGTGGCGGAGTCAATGCTCTAGAGGCAATGA 540  
Qy 541 GCTACGGCGGTGAGAGTGAACCTTAATGGCGCGGAGAGAGAGAGAGAGAGAGAG 600  
Db 541 GCTACGGCGGTGAGAGTGAACCTTAATGGCGCGGAGAGAGAGAGAGAGAGAGAG 600  
Qy 601 TTCTACTCTCAACAG 660  
Db 601 TTCTACTCTCAACAG 660  
Qy 661 TGGCTTCTTCAATGGGTGGTGGGGGGAAGAAGATATCCCTTCTTATGGGAAATCCA 720  
Db 661 TGGCTTCTTCAATGGGTGGTGGGGGGAAGAAGATATCCCTTCTTATGGGAAATCCA 720  
Qy 721 GCTTACTGTGCAACCCAGGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 721 GCTTACTGTGCAACCCAGGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Qy 781 GTAGTGTGCTTGAAG 840  
Db 781 GTAGTGTGCTTGAAG 840  
Qy 841 AAG 900  
Db 841 AAG 900  
Qy 901 AAGCAG 960  
Db 901 AAGCAG 960  
Qy 961 GAGATGATATCATTAATTCGGCTTGTGGGTTCTTCTTCACTACAGAGATGATGC 1020  
Db 961 GAGATGATATCATTAATTCGGCTTGTGGGTTCTTCTTCACTACAGAGATGATGC 1020  
Qy 1021 ATCCGATGATATGATGATTTGCTTGTGGCTATATCTGTATCTGTGAATTTGGTTTG 1080  
Db 1021 ATCCGATGATATGATGATTTGCTTGTGGCTATATCTGTATCTGTGAATTTGGTTTG 1080  
Qy 1081 CCATGCTTGATTTCTTGATCAATTCGAAAGTGTTCCTATTTAGAGAGAGAGAGAG 1140  
Db 1081 CCATGCTTGATTTCTTGATCAATTCGAAAGTGTTCCTATTTAGAGAGAGAGAGAG 1140

Db 1081 CCATGCTTGATTTCTTGATCAATTCGAAAGTGTTCCTATTTAGAGAGAGAGAGAG 1140  
Qy 1141 TAGACCGGCTGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
Db 1141 TAGACCGGCTGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
Qy 1201 ATTTCTTGTGAGAGAGTGTGATCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
Db 1201 ATTTCTTGTGAGAGAGTGTGATCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
Qy 1261 TTCTATCTATCTTTCGGTGGATTAATCTGTGATTAAGTGTCTGTATGTTCTGATG 1320  
Db 1261 TTCTATCTATCTTTCGGTGGATTAATCTGTGATTAAGTGTCTGTATGTTCTGATG 1320  
Qy 1321 ATGGTGTGAG 1380  
Db 1321 ATGGTGTGAG 1380  
Qy 1381 GGGTTCCTTCTGAG 1440  
Db 1381 GGGTTCCTTCTGAG 1440  
Qy 1441 AGAAGATGAGTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Db 1441 AGAAGATGAGTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Qy 1501 TGAAGAGAGAGATGAG 1560  
Db 1501 TGAAGAGAGAGATGAG 1560  
Qy 1561 AAGTTCCTGAG 1620  
Db 1561 AAGTTCCTGAG 1620  
Qy 1621 GTGATCATCTGAGATGATTCAGGTCTTCTTGGCCAAAGGGAGAGAGAGAGAGAG 1680  
Db 1621 GTGATCATCTGAGATGATTCAGGTCTTCTTGGCCAAAGGGAGAGAGAGAGAGAGAG 1680  
Qy 1681 GAAATGAG 1740  
Db 1681 GAAATGAG 1740  
Qy 1741 ATTAAG 1800  
Db 1741 ATTAAG 1800  
Qy 1801 CATATTTGTAACTTGGATTTGTATCACTACATCAACAGAGAGAGAGAGAGAGAGAG 1860  
Db 1801 CATATTTGTAACTTGGATTTGTATCACTACATCAACAGAGAGAGAGAGAGAGAGAG 1860  
Qy 1861 CAATGTGTTTATGATGAG 1920  
Db 1861 CAATGTGTTTATGATGAG 1920  
Qy 1921 AAAGATTTGATGAG 1980  
Db 1921 AAAGATTTGATGAG 1980  
Qy 1981 ATATCAATGAG 2040  
Db 1981 ATATCAATGAG 2040  
Qy 2041 TATTTAG 2100  
Db 2041 TATTTAG 2100  
Qy 2101 GGACTTGAAG 2160  
Db 2101 GGACTTGAAG 2160  
Qy 2161 AAAAG 2220  
Db 2161 AAAAG 2220







DB 2101 GAGCTGGAACGCTGCGCCCAAGTGTGTTTGTGCTGCTTGGCAATGGAAGC 2160  
QY 2161 AAAAGAGACTACCAACCACAAACAGAGAGAAAAAGTTATTTTCAAGAAAG 2220  
DB 2161 AAAAGAGACTACCAACCACAAACAGAGAGAAAAAGTTATTTTCAAGAAAG 2220  
QY 2221 AGAACCAATCCCTGCGATATGCTCTGGGAAATTGAGGAAGCTCTCAGAGAGTGA 2280  
DB 2221 AGAACCAATCCCTGCGATATGCTCTGGGAAATTGAGGAAGCTCTCAGAGAGTGA 2280  
QY 2281 ATGAAAAGCCGCTATTTGTAATCAACAAAATTGAAAAGAAATTTGGCCAACTTCTG 2340  
DB 2281 ATGAAAAGCCGCTATTTGTAATCAACAAAATTGAAAAGAAATTTGGCCAACTTCTG 2340  
QY 2341 TTTTGTATACATCCACACTTCTCGAAGATGTGAACTTGAAGAGTGCAGTCTGCTT 2400  
DB 2341 TTTTGTATACATCCACACTTCTCGAAGATGTGAACTTGAAGAGTGCAGTCTGCTT 2400  
QY 2401 CTCTTTGAAGAAGCTATACATGCTATGTTGGTTATGAAACAAGACAGCTGGG 2460  
DB 2401 CTCTTTGAAGAAGCTATACATGCTATGTTGGTTATGAAACAAGACAGCTGGG 2460  
QY 2461 GAAAAGAGATTGCTGATCTATGATCAGATCAGAAATTTCTAAGTGTTCAGA 2520  
DB 2461 GAAAAGAGATTGCTGATCTATGATCAGATCAGAAATTTCTAAGTGTTCAGA 2520  
QY 2521 TGCATTTGTCATGCTGCGGCTCAATTTTACTGCAATCCTAAACGAGTTGCATTCAAGGTT 2580  
DB 2521 TGCATTTGTCATGCTGCGGCTCAATTTTACTGCAATCCTAAACGAGTTGCATTCAAGGTT 2580  
QY 2581 CTGCACCTCTGATTTTTCAGATGCTCTTACACAGAGTCTTGGTGGCTCTTGGGTCTA 2640  
DB 2581 CTGCACCTCTGATTTTTCAGATGCTCTTACACAGAGTCTTGGTGGCTCTTGGGTCTA 2640  
QY 2641 TTGAGATCTTCTGAGCAATCATGCCCCCTTGGTATGGGTATGGTGGCGGTCTGAAT 2700  
DB 2641 TTGAGATCTTCTGAGCAATCATGCCCCCTTGGTATGGGTATGGTGGCGGTCTGAAT 2700  
QY 2701 TTTTGAAGAATTTTCTCAATCAATCCATGCTGATCTTGAACATCTATCCCTCT 2760  
DB 2701 TTTTGAAGAATTTTCTCAATCAATCCATGCTGATCTTGAACATCTATCCCTCT 2760  
QY 2761 TGGCTTACTGATCATGCTGCGCATCTGTTTATTAACAGGAAATTTATCACTCCAGAGC 2820  
DB 2761 TGGCTTACTGATCATGCTGCGCATCTGTTTATTAACAGGAAATTTATCACTCCAGAGC 2820  
QY 2821 TGAATTAATGTTGCCAGCCTGAGTTCATGTCACCTTTTATCTGCAATTTTGTCTACAGCA 2880  
DB 2821 TGAATTAATGTTGCCAGCCTGAGTTCATGTCACCTTTTATCTGCAATTTTGTCTACAGCA 2880  
QY 2881 TCTTGAAGAATGATGAGTGTGTTGAAATTTGATGACTGTGAGAGATGAGCACTTCT 2940  
DB 2881 TCTTGAAGAATGATGAGTGTGTTGAAATTTGATGACTGTGAGAGATGAGCACTTCT 2940  
QY 2941 GGGTCATTTGAGAGTGTGCTTCAACCTCTCTTGTCTGTGTTCCAGGACTTCTCAAGTCA 3000  
DB 2941 GGGTCATTTGAGAGTGTGCTTCAACCTCTCTTGTCTGTGTTCCAGGACTTCTCAAGTCA 3000  
QY 3001 TAGCTGCGTTGATGATCAAGCTTCAACCGTGAATCAAAAGGTGGAATATAGAGAGTTCT 3060  
DB 3001 TAGCTGCGTTGATGATCAAGCTTCAACCGTGAATCAAAAGGTGGAATATAGAGAGTTCT 3060  
QY 3061 CAGAGCTATATACATTCAATGAGACTACCTTATTTGATACCTCTCAACCTTGTCTTAT 3120  
DB 3061 CAGAGCTATATACATTCAATGAGACTACCTTATTTGATACCTCTCAACCTTGTCTTAT 3120  
QY 3121 TGAATTTATTTGTTGTTGCTGCTGCGCTTTCAATGCAATCAATACGATATGATCAT 3180  
DB 3121 TGAATTTATTTGTTGTTGCTGCTGCGCTTTCAATGCAATCAATACGATATGATCAT 3180  
QY 3181 GGGGCCCCCTCTTTGGGAAGCATTTCTTGAATTTTGGGTATGTCATCTTATCCCT 3240  
DB 3181 GGGGCCCCCTCTTTGGGAAGCATTTCTTGAATTTTGGGTATGTCATCTTATCCCT 3240

DB 3181 GGGGCCCCCTCTTTGGGAAGCATTTCTTGAATTTTGGGTATGTCATCTTATCCCT 3240  
QY 3241 TTCTCAAAAGTTTGGTGGAAAGCAAAAACAGACAACCAAGATTTGATCTGTGCTCA 3300  
DB 3241 TTCTCAAAAGTTTGGTGGAAAGCAAAAACAGACAACCAAGATTTGATCTGTGCTCA 3300  
QY 3301 TTCTGCTGCTTCAATCTTCTGCTCTTGGGTGGGTTGGATGATCTTCTTCTGGAAG 3360  
DB 3301 TTCTGCTGCTTCAATCTTCTGCTCTTGGGTGGGTTGGATGATCTTCTTCTGGAAG 3360  
QY 3361 ATGATGCTGCTTCTTGGAGAGTGTGTTGATGCACTAGATGTCAGTCACTAG 3420  
DB 3361 ATGATGCTGCTTCTTGGAGAGTGTGTTGATGCACTAGATGTCAGTCACTAG 3420  
QY 3421 CTCCCCCAATTCGATATGCTTGAAGATATTTTCTGTGTTTCTCCCATATTCAGTGT 3480  
DB 3421 CTCCCCCAATTCGATATGCTTGAAGATATTTTCTGTGTTTCTCCCATATTCAGTGT 3480  
QY 3481 CTGTAGATTAAGACATGAATAATGTCAGAGTTCTTTTGTATCCATGGTGAACCTTAA 3540  
DB 3481 CTGTAGATTAAGACATGAATAATGTCAGAGTTCTTTTGTATCCATGGTGAACCTTAA 3540  
QY 3541 TATCTGAGATATATCTGCGGAAATGAGAGCTGCGCAATCCTTGTGAGTGGCGG 3600  
DB 3541 TATCTGAGATATATCTGCGGAAATGAGAGCTGCGCAATCCTTGTGAGTGGCGG 3600  
QY 3601 TGAATACAGATATGCAAGTGTGATTTGATGTCAGCATCTTATTAATCTTGTGCAATA 3660  
DB 3601 TGAATACAGATATGCAAGTGTGATTTGATGTCAGCATCTTATTAATCTTGTGCAATA 3660  
QY 3661 TAGATGGCTGAGCGCAACGCAAGATTTTGTATTTCTGCACTGCTCCCGTCAAACT 3720  
DB 3661 TAGATGGCTGAGCGCAACGCAAGATTTTGTATTTCTGCACTGCTCCCGTCAAACT 3720  
QY 3721 TGGTTCTCAATTAAGGAGGAGCAAGATCATCTGCACTGCAAGCAAGCACTGCACTT 3780  
DB 3721 TGGTTCTCAATTAAGGAGGAGCAAGATCATCTGCACTGCAAGCAAGCACTGCACTT 3780  
QY 3781 ATTTATGATGCTGCTTCAATTTGAGGCTTGTTCATTAATCATGTTGCTTATACAGAAA 3840  
DB 3781 ATTTATGATGCTGCTTCAATTTGAGGCTTGTTCATTAATCATGTTGCTTATACAGAAA 3840  
QY 3841 AACGAATATTAGATTAATCTATGATTAATTAAGTATGAATGAGCCCTGTTTGTGT 3900  
DB 3841 AACGAATATTAGATTAATCTATGATTAATTAAGTATGAATGAGCCCTGTTTGTGT 3900  
QY 3901 TGTGTACTGTATCATCTGAGTGTGTTTGTGAATA 3936  
DB 3901 TGTGTACTGTATCATCTGAGTGTGTTTGTGAATA 3936  
  
RESULT 5  
US-09-900-237-3  
; Sequence 3, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 3  
; LENGTH: 3776  
; TYPE: DNA  
; ORGANISM: Zea mays



US-09-900-237-3

Query Match 65.9%; Score 2592.6; DB 10; Length 3776;

Best Local Similarity 86.6%; Pred. No. 0;

Matches 2885; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY 104 CTCTGCTCTTCTGCGCGGCGGGAAGTGGAGGGGAGGAGGAGTGGAGGGCGGCGGGCT 163  
DB 139 GCGTGGCCCCGGGATCTGGAGATCTGTAGCCGACAGGGGAGTGGAGGCGCGCGGGCT 198  
QY 144 GGTGCGCGGCTCCCAACCCGCAACGAGCTCTGTCATCCGCGCGACGCGGATCCCG 223  
DB 199 GGTGCGCGGCTCCCAACCCGCAACGAGCTCTGTCATCCGCGCGACGCGGATCCCG 258  
QY 224 GCGGAGCGCGCGCGGAGCAGAACCGGCGGTGTGCCAGATTGGCGCGACGACGTGCG 283  
DB 259 GCGGAGCGCGCATGGACACAGCGGAAACGCGGATGTGCCAGATTGGCGCGACGACGTGCG 318  
QY 284 CTTTGGCCCCGGGCGGAGACCCCTTGTGGCGGTGGCAACGATGGCGCTTCCCGCTGCGG 343  
DB 319 GCGCAACCCCGACGGGAGCCGTTGCTGGCGGTGGCAACGATGGCGCTTCCCGCTGCGG 378  
QY 344 GGAATGCTACGAAATACGAGCGCGGAGGAGCAGCAGAACTGCGCCCACTGCAAGACTG 403  
DB 379 GGAATGCTACGAAATACGAGCGCGGAGGAGCAGCAGAACTGCGCCCACTGCAAGACTG 438  
QY 404 ATACAGCGCGCTCAAGGGGCTGCAACGTGTGACCGGTGACGAGAGGAGAGCGCGTGA 463  
DB 439 CTTCAAGCGCGCTCAAGGGGCTGCGCGCGGTGCGCGGAGCAGAGGAGAGAGCGCGTGA 498  
QY 464 TGAATGCTACGAAATACGAGCGCGGAGGAGCAGCAGAACTGCGCCCACTGCAAGACTG 520  
DB 499 GCACTGAGAGAGCGAGTTCACTGAGCAGCAGCAAGCAGACTCCCACTGCAAGACTG 558  
QY 521 CATGCTCTACGAGCAGATGAGCTACGCGCGGTGAGCGGAGGAGCTTAATGCGCGCACAAGC 580  
DB 559 CATGCTCTACGAGCAGATGAGCTACGCGCGGTGAGCGGAGGAGCTTAATGCGCGCACAAGC 618  
QY 581 TTTTCAGCTCAACCCCAATGTTTCACTCTCTCAACCAACGCGGCAATGTGTGATGACATCC 640  
DB 619 ATTCAACCCCAATGTTTCACTCTCTCTCAACCAACGAGTGTGTGATGACATCC 678  
QY 641 ACCGAGAGCAGCGCGGTGCTCTTTCATGGGTGTGGGGGAAAGAGATACATCC 700  
DB 679 GCGGAGCAGCAGCGCGGTGCTCTTTCATGGGTGTGGGGGAAAGAGATACATCC 738  
QY 701 CTTTCTTATGAGGATCCAGCTTACCTGTGCAACCCAGTCTATGAGCCCATCCAGGA 760  
DB 739 TCTCCGTAAGCGGATCCCACTTCTGTGCAACCCAGTCTATGAGCCCATCCAGGA 798  
QY 761 TCTTGTGCTATGAGGTATGATGTGTTCTTGGAGGAAACGAGTGGAGATTTGGAGCA 820  
DB 799 TCTTGTGCTATGAGGTATGATGTGTTCTTGGAGGAAACGAGTGGAGATTTGGAGCA 858  
QY 821 GAGACAGAGAGATGACACAGCGGGAATGATGTGTGTGATGATGATGATGATGATGATG 880  
DB 859 GAGACAGAGAGATGACACAGCGGGAATGATGTGTGTGATGATGATGATGATGATGATG 918  
QY 881 TGAATCAACCTAATGATGAGAGCAAGCAACATGTCAGAGAAATTCACATTCATC 940  
DB 919 AGATCAACCTAATGATGAGAGCAAGCAACATGTCAGAGAAATTCACATTCATC 978  
QY 941 AAGCCAGATTTAATCCATATAGATGATATATATATATGCGCTGTGTGTGTGTGTGTGT 1000  
DB 979 AAGCCAGATTTAATCCATATAGATGATATATATATATATGCGCTGTGTGTGTGTGTGT 1038  
QY 1001 CTTTCACTACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1060  
DB 1039 CTTTCACTACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098  
QY 1061 TATCTGTGAATCTGT 1120  
DB 1099 GATCTGTGAATCTGT 1158

QY 1121 TATTGAGAGAGACTTACCTAGACCGCGGTGCTACCTAGAGTTTGCAGAGAGGCGCACCC 1180  
DB 1159 TATGAGAGGAGAACTATCTTGAACCGCGGTGCTACCTAGAGTTTGCAGAGAGGCGCACCC 1218  
QY 1181 ATCTCACTGCTCCAAATTTGATTTCTTTGTGATGATGATGATGATGATGATGATGATGAT 1240  
DB 1219 TTTCACTGCTCCCGCTGTTGATTTCTTTGTGATGATGATGATGATGATGATGATGATGAT 1278  
QY 1241 TTTGCTCAACAAATATCTGTTCTATCTATCTTTTGGTGGATTAATCTGTTGATTAAGT 1300  
DB 1279 ATTGGTCACTGCTATATCTGTTCTATCTATCTTTGGTGGATTAATCTGTTGATTAAGT 1338  
QY 1301 TTTTGTGCTATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360  
DB 1339 TTTATGCTACGCTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398  
QY 1361 ATCTGAATTTTCAAGAAATGAGGTGCTCTTGTGCAACCGTACATATATTAACCTGCGGC 1420  
DB 1399 ATCTGAATTTTCAAGAAATGAGGTGCTCTTGTGCAACCGTACATATATTAACCTGCGGC 1458  
QY 1421 TCCAGAGTGTACTTCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1480  
DB 1459 TCCAGAGTGTACTTCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1518  
QY 1481 TGTTAGGAGAGAGAGCAATGAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1540  
DB 1519 TGTTAGGAGAGAGAGCAATGAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1578  
QY 1541 CTTAGTTTCCAAAGAGGCGGAAAGTTTCTGAGAGAGATGAGCAATGCAAGATGGAACCC 1600  
DB 1579 CTTAGTTTCCAAAGAGGCGGAAAGTTTCTGAGAGAGATGAGCAATGCAAGATGGAACCC 1638  
QY 1601 CTGGCTGGAACCAATGTTGCTGATCATCTGGAATGATGATGATGATGATGATGATGATGAT 1660  
DB 1639 ATGGCCCGGAATATATGCTCGGTATGATCATCTGGAATGATGATGATGATGATGATGATGAT 1698  
QY 1661 CGAGAGCTTGAATGATGAGGAAATGAACTGCAAGATGATGATGATGATGATGATGATGAT 1720  
DB 1699 TGTGTGCTGATGATGATGAGGAAATGAACTGCAAGATGATGATGATGATGATGATGATGAT 1758  
QY 1721 AGCAGAGCTATTAACCATATTAAGAAAGTGTGTGATGATGATGATGATGATGATGATGAT 1780  
DB 1759 AGCAGAGCTATTAACCATATTAAGAAAGTGTGTGATGATGATGATGATGATGATGATGAT 1818  
QY 1781 TGTGTGCTATTAACCATATTAAGAAAGTGTGTGATGATGATGATGATGATGATGATGAT 1840  
DB 1819 TGTGTGCTATTAACCATATTAAGAAAGTGTGTGATGATGATGATGATGATGATGATGAT 1878  
QY 1841 CAGCAAGCTATTAAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1900  
DB 1879 TAGTAAAGCTATTAAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1938  
QY 1901 TTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1960  
DB 1939 TTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1998  
QY 1961 CCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020  
DB 1999 CAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2058  
QY 2021 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2080  
DB 2059 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2118  
QY 2081 AACAAAGAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2140  
DB 2119 AACAAAGAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2178  
QY 2141 CTGCTTTGCAATAGAGAGCAAGAGAGCAATGATGATGATGATGATGATGATGATGAT 2194  
DB 2179 CTGCTTTGCAATAGAGAGCAAGAGAGCAATGATGATGATGATGATGATGATGATGAT 2238



QY 594 CCCATGTTCACTCTCTCAACCAACGCGCAAAATGATGATGATCATCTCCACCGGAGCAGCAC 653  
DB 426 CCCAATGTTCCCTCTCTCAACCAACGAGATGATGATGATCATCTCCCGCGAGCAGCAC 485  
QY 654 GCGCTGTGCTCTCTTCAATGGGTGTGGGGGAAAGAGATATATCCCTTCTTATGG 713  
DB 486 GCGCTGTGCTCTCTTCTGTTGCGGGTGGCGGGGGAAGAGATTAACCTCTCCGTTAGGG 545  
QY 714 GATCCAGCTTACCTGTGCAACCCAGGTCTATGAGCCATCCAAAGATCTGTGATAT 773  
DB 546 GATCCCACTCTCTGTGCAACCGAGGTCTATGAGCCCTTCCAAAGATCTGTGATAT 605  
QY 774 GGGTATGTATGTGTCTTGGAAAGAAACGATGAGAAATTTGAAAGACAGACAGAGAG 833  
DB 606 GGGTATGAGGAGCTAGATGAGAAAGAGATGAGAGCTGGAAGACAGAAAGCAGAGAG 665  
QY 834 ATGCACCAAGCGGGGAATGATGTGTGTGATGATGATGATGATGATGATGATGATGAT 893  
DB 666 ATGCACCAAGCGAGAAAGATGGCGGCGATGATGATGATGATGATGATGATGATGAT 725  
QY 894 ATGATGAGCAAGCAACAACTGTCCAGGAAATTCACCTTCATCAAGCCAGATTAT 953  
DB 726 ATGATGAGCACTGACAGCCATGTTCAGAAAGATCCCGCTTCTTCAAGCCAAATCAAC 785  
QY 954 CCATATAGATGATATATCATTTATCGGCTTGTGGGTTCTTCTTCACTACCA 1013  
DB 786 CCTATAGATGAT 845  
QY 1014 GTATGATGATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
DB 846 GTATGATGATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905  
QY 1074 TGGTTGGCATGCTTGTGATTTCTGATCAATTCCAAAGTGTTCCTTATGAGAGAG 1133  
DB 906 TGGTTGGCATGCTTGTGATTTCTGATCAATTCCAAAGTGTTCCTTATGAGAGAG 965  
QY 1134 ACTTACCTAGACCGGCTGTGATGAGGTTGCAAGAGAGAGAGAGAGAGAGAGAGAG 1193  
DB 966 ACTTACCTAGACCGGCTGTGATGAGGTTGCAAGAGAGAGAGAGAGAGAGAGAGAG 1025  
QY 1194 CCATGATGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253  
DB 1026 CCTGTTGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085  
QY 1254 AATACGTTCTATCTATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1313  
DB 1086 AATACGTTCTATCTATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1145  
QY 1314 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373  
DB 1146 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205  
QY 1374 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433  
DB 1206 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1265  
QY 1434 TTTCAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493  
DB 1266 TTTCAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1325  
QY 1494 AGAGCATGAGAGAGAGATGAGAAATTCAGGTGAGAAATTCAGGTGAGAAATTCAGGT 1553  
DB 1326 AGAGCATGAGAGAGAGATGAGAAATTCAGGTGAGAAATTCAGGTGAGAAATTCAGGT 1385  
QY 1554 GCCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613  
DB 1386 GCCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445  
QY 1614 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
DB 1446 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505

QY 1674 TGTGAGGAAATGAACTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1733  
DB 1506 GTGAAAGAAATGAACTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1565  
QY 1734 AACCATGATGAGAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1793  
DB 1566 AACCATGATGAGAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1625  
QY 1794 AATGCTCATATTTTGTAACTTGTGATGATGATGATGATGATGATGATGATGATGAT 1853  
DB 1626 AATGCTCATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685  
QY 1854 AAGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1913  
DB 1686 AAGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1745  
QY 1914 TTTCCCTCAAGATTTGATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1973  
DB 1746 TTTCCCTCAAGATTTGATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1805  
QY 1974 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033  
DB 1806 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1865  
QY 2034 GATGTGATTTTGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2093  
DB 1866 GATGTGATTTTGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1925  
QY 2094 CCATCAAGGATTTGAACTGCTGCGCCCAAGGTGCTTGTGATGATGATGATGATGAT 2153  
DB 1926 CCATCAAGGATTTGAACTGCTGCGCCCAAGGTGCTTGTGATGATGATGATGATGAT 1985  
QY 2154 ACGAAGCAAAAGAGACTACCAACCCAAACAGAGAGAGAGAGAGAGAGAGAGAGAG 2207  
DB 1986 ACGAAGCAAAAG 2045  
QY 2208 TTCAAGAAAG 2267  
DB 2046 TTCAAGAAAG 2105  
QY 2268 CCAGAGCTGAG 2327  
DB 2106 CCAGAGCTGAG 2165  
QY 2328 GGCATATCTTCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2387  
DB 2166 GGCATATCTTCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2225  
QY 2388 GCAAGTCTGCTTCTTGTGAAAGCTATGATGATGATGATGATGATGATGATGATGAT 2447  
DB 2226 GCAAGTCTGCTTCTTGTGAAAGCTATGATGATGATGATGATGATGATGATGATGAT 2285  
QY 2448 AAGACAGACTGGGAG 2507  
DB 2286 AAGACAGACTGGGAG 2345  
QY 2508 ACTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2567  
DB 2346 ACTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2405  
QY 2628 GCTCTTGGGCTATGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGAT 2687  
DB 2406 GCTCTTGGGCTATGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGAT 2465  
QY 2628 GCTCTTGGGCTATGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGAT 2687  
DB 2466 GCTCTTGGGCTATGAGATCTTCTTCAAGCAATGATGATGATGATGATGATGATGATGAT 2525  
QY 2688 GCGGCTGAGAAATTTTGTGAAAGATTTTCTTCAATCAATCCATGATGATGATGATGAT 2747  
DB 2526 GCGGCTGAGAAATTTTGTGAAAGATTTTCTTCAATCAATCCATGATGATGATGATGATGAT 2585  
QY 2748 TCTATTTCCCTCTTGGCTTACTGTACATTTGCTGCAATCTGTTTATGACAGGAAATTT 2807



Db 846 GTGATGATCCGGTCCCTGATGCAATTTGCTTTATGCTCATATCTGTATCTGTAAATT 905  
Qy 1074 TGGTTGGCCAGTCTTGGATCTTGTATCAATTTCCCAAGTGGTTCCCTATTGAGAGAG 1133  
Db 966 TGGTTGGCAATGCTTGGATCTTGTATCAATTTCCCAAGTGGTTCCCTATTGAGAGAG 965  
Qy 1134 ACTTACCTAGACCGGCTGTCACTGAGTTCCAGAGAGAGAGGCGACCTTCAATTGCT 1193  
Db 966 ACCTATCTTGACCGGCTGTGATTGAGTTGACAGAGAGAGGCACTCTTCTCACTCCGC 1025  
Qy 1194 CCAATGATTTCTTGTGAGTGGTGTGATCCCTTAAAGAAACCTCTTGGTCAACA 1253  
Db 1026 CCTGTGATTTCTTGTGAGTGGTGTGATCCCTTAAAGAAACCTCTTGGTCACTGCT 1085  
Qy 1254 AATACGTCTTCTATCTATCTTCCGTGAGTTATCTCTGTGATGAGTTCTTGTGATGTT 1313  
Db 1086 AATACGTCTTCTATCTATCTTCCGTGAGTTATCTCTGTGATGAGTTCTTGTGATGTT 1145  
Qy 1314 TCTGATGATGCTGCTGCAATGCTAAAGTTGAGAGATTAATCTGTTGATGAGTTCTTGTGATGTT 1373  
Db 1146 TCTGATGATGCTGCTGCAATGCTAAAGTTGAGAGATTAATCTGTTGATGAGTTCTTGTGATGTT 1205  
Qy 1374 AAGAAATGGTCTCTTCTGCAAAAGGTACATATTTGAACCTCGGCTCCAGAGTGTAC 1433  
Db 1206 AAGAAATGGTCTCTTCTGCAAAAGGTATGACCTTGAAGCTCGGCTCCAGAGTGTAC 1265  
Qy 1434 TTCCAACAGAGATGACTTCTTGAAGAGAGAGTGGCAGCAACTTTGTTAGAGAGAG 1493  
Db 1266 TTCCAACAGAGATGACTTCTTGAAGAGAGAGTGGCAGCAACTTTGTTAGAGAGAG 1325  
Qy 1494 AGAGCAATGAGAGAGATGATGAGAAATCAAGTGAAGATCAATGCTTACTTGTCCAA 1553  
Db 1326 AGAGCAATGAGAGAGATGATGAGAAATCAAGTGAAGATCAATGCTTACTTGTCCAA 1385  
Qy 1554 GCGCAGAAAGTCCAG 1613  
Db 1386 GCGCAGAAAGTCCAG 1445  
Qy 1614 AATGTTGCTGATCATCTGGAATGATTCAGGTCTTCTTGGCAGAGAGAGAGAGAG 1673  
Db 1446 AATGTTGCTGATCATCTGGAATGATTCAGGTCTTCTTGGCAGAGAGAGAGAGAG 1505  
Qy 1674 TGTGAGGAAATGAATGCTGCAAGATTTGTTTCTTGAAGAGAGAGAGAGAGAGAG 1733  
Db 1506 GTGAGAGAAATGAGCTGCTGATGATGATTTGTTTCTTGAAGAGAGAGAGAGAG 1565  
Qy 1734 AACCATCAATGAAG 1793  
Db 1566 AACCATCAATGAAG 1625  
Qy 1794 AATGCTCCATATTTGTTAACTTGGATTTGATTCATCATCAAGAGAGAGAGAGAG 1853  
Db 1626 AATGCTCCATATTTGTTAACTTGGATTTGATTCATCATCAAGAGAGAGAGAGAG 1685  
Qy 1854 AAG 1913  
Db 1686 AAG 1745  
Qy 1914 TTCCCTCAAG 1973  
Db 1746 TTCCCTCAAG 1805  
Qy 1974 TTTTGTGATATCAACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2033  
Db 1806 TTTTGTGATATCAACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1865  
Qy 2034 GAGATGATATTTGAAG 2093  
Db 1866 GAGATGATATTTGAAG 1925  
Qy 2094 CCAATCAAG 2153  
Db 1926 CCAATCAAG 1985

Qy 2154 AG 2207  
Db 1986 AG 2045  
Qy 2208 TTTCAAG 2267  
Db 2046 TTTCAAG 2105  
Qy 2268 CCAG 2327  
Db 2106 CCAG 2165  
Qy 2328 GCGCAATCTTCTGTTTGTGTTTCAATCAACTTCTGAGAGAGAGAGAGAGAGAGAG 2387  
Db 2166 GCGCAATCTTCTGTTTGTGTTTCAATCAACTTCTGAGAGAGAGAGAGAGAGAGAG 2225  
Qy 2388 GCAAGTCCGCTCTCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2447  
Db 2226 GCAAGTCCGCTCTCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2285  
Qy 2448 AAG 2507  
Db 2286 AAG 2345  
Qy 2508 ACTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2567  
Db 2346 ACTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2405  
Qy 2568 GCAATCAAG 2627  
Db 2406 GCAATCAAG 2465  
Qy 2628 GCTCTGAGTCTATGAGATCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2687  
Db 2466 GCTCTGAGTCTATGAGATCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2525  
Qy 2688 GCGGCTGGAATTTTGGAG 2747  
Db 2526 GCGGCTGGAATTTTGGAG 2585  
Qy 2748 TCTATTCCTCTTGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2807  
Db 2586 TCTATTCCTCTTGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2645  
Qy 2808 ATCACTCAG 2867  
Db 2646 ATCACTCAG 2705  
Qy 2868 TTTGCTAG 2927  
Db 2706 TTTGCTAG 2765  
Qy 2928 AATGAG 2987  
Db 2766 AATGAG 2825  
Qy 2988 CTTCTCAAG 3047  
Db 2826 CTTCTCAAG 2885  
Qy 3048 GATGAG 3107  
Db 2886 GATGAG 2945  
Qy 3108 AACTGCTCTTATGAG 3167  
Db 2946 AACTGCTCTTATGAG 3005  
Qy 3168 GATATGAG 3227  
Db 3006 GATATGAG 3065



Db 1685 GATGGAAGCCCTGGCTGGAACAACGATACCGCATCATCTCGAATATATTCAGGTATTC 1744  
Qy 1650 CTTGGCCAAAGCGAGGCTTGACTGTGAGGGAAATGAATGCCAGATGGTTATGTT 1709  
Db 1745 CTTGGCCAAAGGCGGTGCTGTGAGGAAATGATGGCTCGCTGGTTATATGTC 1804  
Qy 1710 TCTAGAGAAACGACGAGCTATTAACCATTAAGAAAGCTGTGCTATGATGATG 1765  
Db 1805 TCGAAGAAAGAGGCGGTATTAACCATCAAGAAAGCGTGGGCAATGATGACATG 1864  
Qy 1770 GTCCAGTCTCTGCTGTAATAAATGCTCATATTTGTTAACTTGGATTTGATGAC 1829  
Db 1865 GTCCGTGCTCTGCTGTCTTATCAAAATGCTGATACCTATTTGAATTTGATGATGAC 1924  
Qy 1830 TACATCAACAACGAGGCTATTAAGAAAGCAATGTGTTTATGATGAGCCCTTTACTA 1889  
Db 1925 TACATCAACAATAGCAAGGCCATAAAGAGCTATGTGTTTCATGATGATCTTTGGTG 1984  
Qy 1890 GGAAGAAAGCTTGTCTATGATGATGATTCCTCAAAATTTGATGGATTTGATGAC 1949  
Db 1985 GGAAGAAAGTGTGCTATGATGATGATTCCTCAAGGTTGATGATGACAAATATAT 2044  
Qy 1950 CGATATGCTAACCGGAATGTTGCTTTTGTATATCAATGAAAGTTTGGATGATAT 2009  
Db 2045 CGATATGCTAACGAGGAGTGTCTTTTGTATCAATCAATGAAAGTTTGGACGTAAT 2104  
Qy 2010 CAGGTCCTCAATTTATGTTGTTACTGAGATGATTTAGAAAGCAGGACCTTATATGTTAT 2069  
Db 2105 CAGGACCCATTTATGTTGTTACTGAGATGATTTTCAAGCAGGACCTGATATGTTAT 2164  
Qy 2070 GATGCCCCCAAAACAAAGAGCCACCATCAAGGACCTTGCATGCTGGCCCAAGTGTC 2129  
Db 2165 GATGTCCTTAAACAAAGAGCCACCATCAAGGACCTTGCATGCTGGCCCAAGTGTC 2224  
Qy 2110 TTTGCTGTGCTGCTTTGGCAATAGAAAGCAAAAGACTACCAACCCCAACAGAG 2189  
Db 2225 CTCTCTGTGCTGAGCAAGAAAGAAATTAAGAAAGAACTTCAAAACCAAGAGCGAG 2284  
Qy 2190 AAGAAAGATTAATTTTTCAGAAAGAAAGCAATCCCTGCAATGCTCTTGGT 2249  
Db 2285 AAGAAAGATTAATTTTTCAGAAAGCAAAACCATCTCTGCAATGCTCTTGGGT 2344  
Qy 2250 GAAATTTGAGAGCTGCTCAGAGAGTGAATGAAGAGCGCGGATTTGATTAATCA 2309  
Db 2345 GAAATTTGAGAGCTGCTCAGAGAGTGAATGAAGAGCGCGGATTTGATTAATCA 2404  
Qy 2310 AAATTTGAGAGAGATTTGGCCAAATCTTCTGTTTGTATCACTCAACTCTCGAGAT 2369  
Db 2405 AAATTTGAGAGAGATTTGGCCAAATCTTCTGTTTGTATCACTCAACTCTCGAGAT 2464  
Qy 2370 GGTGAGACCTTGAAGAGTGCAGATCTGCTTCTTTTGAAGAGATATACATGCTAT 2429  
Db 2465 GGAGGAGACCTTGAAGAGGCAAGTGCAGATCTTCTTGAAGAGATATATATATC 2524  
Qy 2430 AGTGTGTTATGAAGACAGACAGCTGGGAAAAAGATTTGCTGATCTAGATCA 2489  
Db 2525 AGCTCGGCTCAGAGACAGACAGCTGGGAAAAAGATTTGCTGATTTACGAGATCG 2584  
Qy 2490 GTTACAGAGATATTTCTAATGCTGTTCAAGATGCAATGCTGGGCGCATTTAC 2549  
Db 2585 ATCAAGAGAGATATTTCTAATGCTGTTCAAGATGCAATGCTGGGCGCATTTAC 2644  
Qy 2550 TGCATACCTTAAACGGGTTGATTTCAAGAGTTCTGCACTGATCTTTCAAGATGCT 2609  
Db 2645 TGCATACCGGAACGGGCTGATTTCAAGAGTTCTGCGGCTCTGAACCTTTCCGACGTC 2704  
Qy 2610 CACGAGGCTTGGTGGGCTCTTGGGCTTATGAGATCTTTCAAGCAATATGCTCCT 2669  
Db 2705 CACGAGGCTTGGTGGGCTCTTGGGCTCTGCGATTTCTTCAAGCAAGACCTGCCA 2764  
Qy 2670 CTTTGGATGGTATGATGATGATGATTTTGAAGATTTTCCATCACTCACTCC 2729

Db 2765 CTTTGGATGAGATACGGGCGGCGCTAAATTCCTGAAAGCTTTCTTATATCACTCC 2824  
Qy 2730 ATGATGATATCTTGGACATCTATTCCTCTTGGCTTACTGTATATGCTCCATGCTGT 2789  
Db 2825 ATGATGATATCTTGGACATCTATTCCTCTTGGCTTACTGTATATGCTCCATGCTGT 2884  
Qy 2790 TTATTTGAGAGGAAATTTATCACTCCAGAGCTGAATATATGTCACCTGTGTTATG 2849  
Db 2885 CTGCTACGGGAGGATTTATCAACAGAGCTTACCAATGTCGCAATATCTGATGATG 2944  
Qy 2850 TCACTTTTATCTGATTTTGTCTACAGAGATCTTAAGAAATGAGATGAGTGTGTTGA 2909  
Db 2945 GCACTTTTATCTGATCTCCGTGACCGGATCTTGAATGAGTGTGAGTGTGTTG 3004  
Qy 2910 ATGATGATCTGCTGAGGAAATGAGAGCTTGGGCTATTTGAGAGTGTGCTCAACCTC 2969  
Db 3005 ATCAAGCACTGCTGAGGAAAGAGAGCTTGGGCTATCGAGAGCGTTTGGCGCATCTG 3064  
Qy 2970 TTTGCTGTGTTCCAGGACCTTCAAGGTGATGATGCTGTGTTATATCAAGCTTCAACCTG 3029  
Db 3065 TTGCGGTGTGTCAGGCGCTGCTGAAAGGTGTGCGCGCAATCAAGAGCTTCAACCTG 3124  
Qy 3030 ACATCAAGGCTGAGATGATGAGAGGATTCAGAGCTATATATCAATCAATGAGCTTACC 3089  
Db 3125 ACATCAAGGCTGAGAGGAGAGAGGATTCAGAGCTATATATCAATCAATGAGCTTACC 3184  
Qy 3090 TTATTTATCTCTTACCACTTGTCTTATTTGAATCTTATGATGCTGTGCTGAGCT 3149  
Db 3185 CTGCTGATATCTTACCACTTGTCTTGTGATTTCAATGAGGAGTGTGCTGAGCT 3244  
Qy 3150 TCAATGCTGATATCAATGAGATGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGG 3209  
Db 3245 TCGAAGCGATGAAACAGGATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3304  
Qy 3210 GCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3269  
Db 3305 GCCTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3364  
Qy 3270 AGAAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3329  
Db 3365 AGAAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3424  
Qy 3330 TGGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3389  
Db 3425 TGGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3484  
Qy 3390 TTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3402  
Db 3485 CTGATCTGCACT 3497

RESULT 9  
US-10-209-059-17  
; Sequence 17, Application US/10209059  
; Publication No. US2003016383A1  
; GENERAL INFORMATION:  
; APPLICANT: Dhugra, Kanwarpal S.  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; FILE REFERENCE: 0864R2  
; CURRENT APPLICATION NUMBER: US/10/209,059  
; PRIOR FILING DATE: 2002-07-31  
; PRIOR APPLICATION NUMBER: 60/096,822  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 09/371,383  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: 09/550,483  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 3813





Db 2225 CTCCTTGCTGCTGACGAGAACAGATTAAGAAAGACTTACAAAACCAAGACGAG 2284  
Qy 2190 AAGAAAAAGTTATATTTTCAAGAAAGAGAAACAATCCCTGCATATGCTTGCT 2249  
Db 2285 AAGAGAAAAAGATTTATTTTCAAGAAAGAGAAACCAATCTCTGCATATATGCTTG 2344  
Qy 2250 GAAATTTGAGAGAGCTGCTCCAGAGAGTGAATGAAGAGCCGGATTTGTAAATCAACA 2309  
Db 2345 GAAATTTGAGAGAGCTGCTCCAGAGAGTGAATGAAGAGCCGGATTTGTAAATCAACA 2404  
Qy 2310 AAAATTAGAAAAAGATTTTGCCCAATCTTCTGTTTTTGTATACATCAACCTTCCAGAT 2369  
Db 2405 AAACTAGAGAGAAATTTGGGAGATCTTCTGTTTTTGTGATCAACACTTTTAGAAG 2464  
Qy 2370 GGTGAAACCTTGAAGAGTGAAGTCTGCTTCTTTTGAAGAAAGTATATATGATTCAT 2429  
Db 2465 GGAGGAGACCTTGAAAGGCGCAAGTCCAGCTTCTCTTGAAGAAAGTATATATGATTCAT 2524  
Qy 2430 AGTTGTGTTTGAAGACAAACAGATGCGGAAAAAGATTTGGCTGATCTATGATCA 2489  
Db 2525 AGCTGCGCTACGAAAGACAAAGCCAGCTGGGAAAAAGATTTGGCTGATTTACGATCG 2584  
Qy 2490 GTTACAGAGATATTTCTAATGATTTCAAGATGCAATGATGATTTGGCGGCTCAAT 2549  
Db 2585 ATCAAGAGAGATATCTTGACTGATTTAAGATGCACTGCTGCTGCGGCTATTTTAC 2644  
Qy 2550 TGCAATACCTAAACGGGTTGCAATTAAGGTTCTGACCTCTGATCTTTTCAGATGCT 2609  
Db 2645 TGCAATCCGAAACGGGCTGCTCAATTAAGGTTCTGCGGCTCTGAACTTTCCGACGCT 2704  
Qy 2610 CACCAAGTCTTCCGTTGGGCTCTTGGGCTTATGAGATTTTCTTCAAGCATATTTGCC 2669  
Db 2705 CACCAAGTCTTCCGTTGGGCTCTTGGGCTTATGAGATTTTCTTCAAGCATATTTGCC 2764  
Qy 2670 CTTTGGTATGGGATATGATGATGATGATGATTTTGAAGAAATTTTCCATCAATCACT 2729  
Db 2765 CTTTGGTATGATATGATGATGATGATGATTTTGAAGAAATTTTCTTATCAATCTC 2824  
Qy 2730 ATCGTATATCTTGAAGATCTATTCCTCTTGGCTTACATGCTGCTGCTCAATCTGT 2789  
Db 2825 ATCGTATATCTTGAAGATCTATTCCTCTTGGCTTACATGCTGCTGCTCAATCTGT 2884  
Qy 2790 TTATGACAGGAAATTTATCTCAAGCTGATTAATGTTGCCAGCTGTTGTTCAATG 2849  
Db 2885 CTGCTCACGGGAAAGTTATCAACACAGGCTTACCAATGTCGCAAGTATCTGTTCAATG 2944  
Qy 2850 TCACTTTTATCTGCAATTTTGTCTAGACATCTTGAAGATGAGATGATGATGATG 2909  
Db 2945 GCACTTTTATCTGCAATCTTGTGACCGGATCTTGAAGATGAGATGATGATGATG 3004  
Qy 2910 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2969  
Db 3005 ATCGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3064  
Qy 2970 TTTGCTGTTTCAAGGATTTTCAAGGCTATGCTGATGATGATGATGATGATGATGATG 3029  
Db 3065 TTTGCGGTTTCAAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3124  
Qy 3030 ACATCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3089  
Db 3125 AGCTGAGAGGCGGAGGAGCAAGAGATTTCTGAGATGATGATGATGATGATGATGATG 3184  
Qy 3090 TTATGATATCTTCAACACTTGTCTTATGAACTTCAATGATGATGATGATGATGATG 3149  
Db 3185 CTGCTGATATCCCGGACAGGCTCTCTGCTGAACTTCAATGAGGATGATGATGATGATG 3244  
Qy 3150 TCAATGCTGATCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3209  
Db 3245 TCGAAGCTGATCAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3304  
Qy 3210 GCAATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3269  
Db 3305 GCTTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3364

Qy 3270 AGAACCACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3329  
Db 3365 AGAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3424  
Qy 3330 TGGGTTGGATTTGATCTTTCTTCTGCGAAGATGATGATGATGATGATGATGATGATGATG 3389  
Db 3425 TGGGTTGGATTTGATCTTTCTTCTGCGAAGATGATGATGATGATGATGATGATGATGATG 3484  
Qy 3390 TTGATGCACT 3402  
Db 3485 CTGATGCACT 3497

RESULT 10  
US-10-160-719-25  
; Sequence 25, Application US/10160719  
; Publication No US20030167528A1  
; GENERAL INFORMATION:  
; APPLICANT: Kanwarpal S. Dhugra  
; APPLICANT: Timothy G. Helentjaris  
; APPLICANT: Benjamin A. Bowen  
; APPLICANT: Xun Mang  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 0864  
; CURRENT APPLICATION NUMBER: US/10/160, 719  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/371, 383A  
; PRIOR FILING DATE: 1998-08-06  
; PRIOR APPLICATION NUMBER: US 60/096, 822  
; PRIOR FILING DATE: 1998-08-17  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 3813  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (215)...(3496)  
US-10-160-719-25

Query Match 56.9%; Score 2241; DB 13; Length 3813;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

Qy 123 GGAAGTGAAGGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 182  
Db 194 GGGCGGGGCGGGGCGCTGAGATGAGAGCTAGCCGGGGCTGATGAGATGAGATGAGATGAG 253  
Qy 183 CGCAACGAGCTGCTGCTATCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241  
Db 254 CGGAACGAGCTGCTGCTATCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313  
Qy 242 GCAGAACGGAGC-----AGTGTGCGAGATTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
Db 314 GCGGCGGGGCGAGAGCGCGCTGCGAGATATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373  
Qy 297 GGGAGACCTTCTGCTGCTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 356  
Db 374 GGGAGACCTTCTGCTGCTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 433  
Qy 357 TACGAGCGCGGAG 416  
Db 434 TACGAGCGCGGAG 493  
Qy 417 AAGGCTGCGAAGCTGATGAG 476  
Db 494 AAGGCTGCGCGCGGATGAG 553  
Qy 477 GAGTTCAACTGG-----GAGGCGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 521

Db 554 GAGTTTCGGCTGCAGAGCGGCGCGCCACAGAGACGCCGAGTACGTGCGGAGTCC 613  
Qy 522 ATGCTTAAGCGCCACATAGACTACGCCGTGGAGGTACCTTAATGGCGGCCCAAGCT 581  
Db 614 ATGCTTAAGCGCGAGATGAGCTACGCGCGCGCGCA-----CGCGCACCGCGC 664  
Qy 582 TTCCAGCTCAACCCCAATGTTCCACTCCTCAACCAACGGCGCAATGTGAAATGATCCCA 641  
Db 665 TTCAGCCCGGTCCTCAACGTCGCTCTCAACCAACGGCGCAATGTGAAATGATCCCG 724  
Qy 642 CCGAGAGACAGCGCGCTGTGCTCTTCTTTCATG-----GGTGGTGGGGGAAAAGG 692  
Db 725 CCGAGAGACAGCGCGCTGTGCTCTTCTTTCATGAGCGCGCGCGCGCGGGGGGAAAGAG 784  
Qy 693 ATACATCCCTCTCTTATGCGGATCCCAAGCTTACCTGTGCAACCCAGGCTATGAGACCA 752  
Db 785 ATCCACCGCTCCCTTTCGAGATCCCACTTCAAGTCAACCGAGATTCATGACCGG 844  
Qy 753 TTCGAAGATCTTGTGCAATAGGGTATGTTGTGTTGGAGGAACGATGGAAT 812  
Db 845 TTCGAAGATCTTGTGCAATAGGGTATGTTGTGTTGGAGGAAGATGGAAGGC 904  
Qy 813 TGAAGACAGACAAAGAGATGACCAAGCGGGAAATGATGGTGGTGGAT--GAT 869  
Db 905 TGAAGACAGACAGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 964  
Qy 870 GGTACGATGCTGATCTACCACTAATGATGAGCAACAACTGTCCAGGAAAT 929  
Db 965 GGCACGATGAGATCTGCTCACTAATGATGAGCAACAACTGTCCAGGAAATGTC 1024  
Qy 930 CCATCTTCATTAACCGCAATTAATCCATATGATGATTAATCATTAATTCGCTGTGCT 989  
Db 1025 CCAATATCATCAACCGCAATTAATCCATAGATGATTAATCATTAATTCGCTGTGCT 1084  
Qy 990 TTGGAGTTCTCTTCTCACTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1049  
Db 1085 TTGGAGTTCTCTTCTCACTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1144  
Qy 1050 CTCAATCTGTTATCTGTAATCTGTTGCAATGTTGCAATGTTGCAATGTTGCAATGTT 1109  
Db 1145 CTCAATCTGTTATCTGTAATCTGTTGCAATGTTGCAATGTTGCAATGTTGCAATGTT 1204  
Qy 1110 AAGGTGTTCCCTATGTAAGAGAGACTTAACCGGCTGTGCTGTGCTGTGCTGTGCTGTG 1169  
Db 1205 AAGGTGCTTCAATGAGAGAGAGACTTAACCGGCTGTGCTGTGCTGTGCTGTGCTGTG 1264  
Qy 1170 GAAGGACGAGCATGCACTGCTCAATGATTTCTTGTGAGTGGTGTGATCCCTTA 1229  
Db 1265 GAAGGTCACCCCTCAAGCTGTGCTCAATGATTTCTTGTGAGTGGTGTGATCCCA 1324  
Qy 1230 AAGAACCTCTTGTGTAACAACAATCTGTTCTATCTATCTTTCGTTGAGTATCT 1289  
Db 1325 AAGAACCTCTTGTGTAACAACAATCTGTTCTATCTATCTTTCGTTGAGTATCT 1384  
Qy 1290 GTTGAATAGGTTCTTGTGATGTTCTGATGATGATGATGATGATGATGATGATGATGAT 1349  
Db 1385 GTTGAATAGGTTCTTGTGATGTTCTGATGATGATGATGATGATGATGATGATGATGAT 1444  
Qy 1350 TTATCTGAACATGTAATTTGCAAAAGATGGGTTCTTCTGCAAAACGTTAATAT 1409  
Db 1445 TTGCTGAAACATGTAATTTGCAAAAGATGGGTTCTTCTGCAAAACGTTAATAT 1504  
Qy 1410 GAACCTCGGCTCCAGAGTGTACTTCAACAGAAATAGACTCTTGAAGACAAAGTG 1469  
Db 1505 GAGCTCGTGTCTCGAGTGTGTAATCCAAAGAAATAGACTCTTGAAGACAAAGTG 1564  
Qy 1470 GCAGCAAACTTTGTAAGAGAGAGAGCAATGAAGAGAGATGAAGAAATTCAGAGTG 1529  
Db 1565 GCTCTTCAATTTGTAAGAGAGAGAGCGGCGATGAAGAGAAATCGAGAAATTCAGAGTA 1624  
Qy 1530 AGAATCAATGCTTATGTTGCCAAGCCAGAAAGTTCTTGAAGAGATGGAATGCAATGCA 1589  
Db 1625 AGAATCAATGCTTATGTTGCCAAGCCAGAAAGTTCTTGAAGAGATGGAATGCAATGCA 1684

Qy 1590 GATGAACCCCTGCGCTGGAACCAATGTTGATGATCTGGAATGATTCAGGTTTC 1649  
Db 1685 GATGAAGCCCTGCGCTGGAACCAATGATGATCTGGAATGATTCAGGTTTC 1744  
Qy 1650 CTTGGCCAAAGCGGAGGCTTGACTGTGAGGGGAAATGAATGCAAGATGTTTATGTT 1709  
Db 1745 CTTGGCCAAAGCGGAGGCTTGAATGAGGAAATGAATGAGTGGCTGCGCTGTTTATGTC 1804  
Qy 1710 TCTAGAGAAAGACAGAGCTATTAACCATCATTAAGAAAGCTGATGATGAATGCAATG 1769  
Db 1805 TCGAGAGAAAGAGGCGAGGTATTAACCATCAAGAAAGCTGATGATGAATGCAATG 1864  
Qy 1770 GTCCAGTCTCTGCTGATCAACAAATGCTCAATTTGTTAACTTGGATGATGATC 1829  
Db 1865 GTCCAGTCTCTGCTGATCAACAAATGCTCAATTTGTTAACTTGGATGATGATC 1924  
Qy 1830 TACATCAACAAAGCAAGCTATTAAGAAAGCAATGTTTATGATGAGCCCTTATCTA 1889  
Db 1925 TACATCAACAAAGCAAGCTATTAAGAAAGCAATGTTTATGATGAGCCCTTATCTA 1984  
Qy 1890 GGAAGAGAGTTGCTATGATGACAGTTCCCTCAAGATTTGATGAGATGATGATGATGAT 1949  
Db 1985 GGAAGAGAGTTGCTATGATGACAGTTCCCTCAAGATTTGATGATGATGATGATGATGAT 2044  
Qy 1950 GATATGCTAACCGGAAATGTTGCTTTTGTATCAACATGAAGAGTTGATGATGAT 2009  
Db 2045 GATATGCTAACCGGAAATGTTGCTTTTGTATCAACATGAAGAGTTGATGATGAT 2104  
Qy 2010 CAGGTCCTAATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2069  
Db 2105 CAGGTCCTAATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2164  
Qy 2070 GATGCCCCCAAAACAAAGAGCCACATCAAGACTTGCATGCTGCGCCAAAGTGTG 2129  
Db 2165 GATGCCCCCAAAACAAAGAGCCACATCAAGACTTGCATGCTGCGCCAAAGTGTG 2224  
Qy 2130 TTTGCTGTTGCTGCTTTGGCAATAGAGCAAAAGAGCTAACAAACCAAAACAGAG 2189  
Db 2225 TTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2284  
Qy 2190 AAGAAAGATTAATTTTCAAGAAAGAGAAACCAATCCCTGATGATGCTGCTGCT 2249  
Db 2285 AAGAAAGATTAATTTTCAAGAAAGAGAAACCAATCCCTGATGATGCTGCTGCT 2304  
Qy 2250 GAAATGACAGAGCTGCTCAGAGCTGAGATGAGAAAGGCGGATATGTAATCAACAA 2309  
Db 2345 GAAATGATGAGAGCTGCTCAGAGCTGAGATGAGAAAGGCGGATATGTAATCAACAA 2404  
Qy 2310 AAATTAAGAAAGAAATTTGCAATCTTCTGTTTGTATCAATCAACTTTCGAGAT 2369  
Db 2405 AAATTAAGAAAGAAATTTGCAATCTTCTGTTTGTATCAATCAACTTTCGAGAT 2464  
Qy 2370 GGTGGAACCTTGAAGAGTCAAGTCTGCTGCTTCTTGAAGAGATTAATCAATGAT 2429  
Db 2465 GGTGGAACCTTGAAGAGTCAAGTCTGCTGCTTCTTGAAGAGATTAATCAATGAT 2524  
Qy 2430 AGTTGATGATGAG 2489  
Db 2525 AGTTGATGATGAG 2584  
Qy 2490 GTTACAGAGATATTTCAATCTGTTTCAAGATCATTTGATGATGATGATGATGATGAT 2549  
Db 2585 GTTACAGAGATATTTCAATCTGTTTCAAGATCATTTGATGATGATGATGATGATGAT 2604  
Qy 2550 TGAATCTTAAGAGGTTGATCAAAAGTTTTCAGACTTCTTCAAGTCTTCTTCAAGTCT 2609  
Db 2645 TGAATCTTAAGAGGTTGATCAAAAGTTTTCAGACTTCTTCAAGTCTTCTTCAAGTCT 2704  
Qy 2610 CACAGAGTCTTCTGAGAGGCTTGGGCTCAATGAGATCTTCTCAAGATCAATGAGCT 2669  
Db 2705 CACAGAGTCTTCTGAGAGGCTTGGGCTCAATGAGATCTTCTCAAGATCAATGAGCT 2764

Qy	2670	TTTTGGTATGGGTATGGGTGGGGGCTGAAATTTTGGAAATATTTCTCATCAACCTC	2729
Db	2765	CTTTGGTACGGATACGGCGGGGGCTAAATTTCTCGAAAGATTTCTTATATCACTCC	2824
Qy	2730	ATCGTGTATCCCTTGACATCTATTTCCCTCTTGGCTTACTGTACATGCTCCGATCTGT	2789
Db	2825	ATCGTTATCCCTGGAGCTCCATCTCTCTCCGGCTTACTGTACTCTTCCGCACTCGC	2884
Qy	2790	TTTATTGACAGGAAATTTTATCATCTCCAGAGCTGAATTAATGTTCCAGCCTGTGTTCA	2849
Db	2885	CTGCTCAACGGGAAATTATATCAACACAAAGCTTACCAATATGTCGCGACATATCTGTTCA	2944
Qy	2850	TCACCTTTTATCTGTGATTTTGGTCTACAGAGCATCCGAAAGATGAGATGGGTGTTGGA	2909
Db	2945	GCACTTTTCACTCTGTGATCTCCGTGACGGACATCTCGGAAATGAGGTGAGTGGCGTGGCC	3004
Qy	2910	ATTGATGACTGTGTGAGAGATAGACAGATTTCTGGGTCAATTGAGAGTGTCTCTCAACCTC	2969
Db	3005	ATCGACGACTGTGTGAGAGAACGAGCATGTTCTGGGTCAATCGGAGGGTTTCGGCGCATGTG	3064
Qy	2970	TTTGTCTGTTTCCAGGACCTTCTCAAGTCAATAGCTGTTGATATACAGCTTCAACCGTG	3029
Db	3065	TTCCGGGTGTTCCAGGGGCTGTGTAAGGTGTTCGCCGGCATCGACAGACTTCAACCGTG	3124
Qy	3030	ACATCAAAAGGTTGAGATGATGAGAGATTTCTCAAGCTATATATCACTTCAATATGACATAC	3089
Db	3125	ACGTGGAAGGCGGGGAGACGAGAGATTTCTCGAGCTGTACAGTTCATAGTGAACCAAC	3184
Qy	3090	TTATTGATACCTCCATACCAACCTGTGTTTATTGAACTTCAATGGTGTGGTGGCGGTT	3149
Db	3185	CTGCTGATATCCCGGACCAACGCTCTCTCTGTAATCTTCAATCGGGTGTGTGGCCGGATC	3244
Qy	3150	TCAATGCGATCAATTAACGATATAGTATCATGAGGCCCCCTTTTGGAAAGCTATCTTT	3209
Db	3245	TCGAACGGATCAACCAACGGGATGAGATGATGAGGGCCCCCTGTTGGGAAAGCTCTTCTC	3304
Qy	3210	GCATTTTGGGTGATGTGTCACTTTATCCCTTTTCTCAAGTTTGTGTGGAAGCAATAC	3269
Db	3305	GCTTCTGGGTGATGTGTCACTGTACCTGTACCCGTCTTCTCAAGGTGTGTGGAGGACAGAAC	3364
Qy	3270	AGGACACCAACGATTTGATCGTGTGTCATTCCTGCTGGGCTCAATCTTCTCGCTCTT	3329
Db	3365	AGGAGCCCGACGATGTGTATGTCTGTGTCACTCTGCTGGCTCGATTTCTCGTCTCTG	3424
Qy	3330	TGGGTTCCGATTTGATCTTTCTTTCTTTCGGAAGATGATGTCCGCTTCTTGAAGAATGTGT	3389
Db	3425	TGGGTCCCGGTGACCCGTTCTCTGCGCAAGACAAAGCCCGCTCTCTGAGAGATGTGGC	3484
Qy	3390	TTGGATTGCAACT 3402	
Db	3485	CTGGACTGCAACT 3497	
RESULT 11			
US-10-160-719-45			
Sequence 45, Application US/10160719			
Publication No. US20030167528A1			
GENERAL INFORMATION:			
APPLICANT: Kanwarpal S. Dhugra			
APPLICANT: Timothy G. Helentjaris			
APPLICANT: Benjamin A. Bowen			
APPLICANT: Xun Wang			
TITLE OF INVENTION: Maize Cellulose Syntheses and Uses			
TITLE OF INVENTION: Thereof			
FILE REFERENCE: 0864			
CURRENT APPLICATION NUMBER: US/10/160,719			
CURRENT FILING DATE: 2002-06-03			
PRIOR APPLICATION NUMBER: US/09/371,383A			
PRIOR FILING DATE: 1999-08-06			
PRIOR APPLICATION NUMBER: US 60/096,822			
PRIOR FILING DATE: 1998-08-17			
NUMBER OF SEQ ID NOS: 60			
SOFTWARE: FastSeq for Windows Version 3.0			

```

; SEQ ID NO 45
; LENGTH: 3813
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215) ... (3496)
US-10-160-719-45

Query Match      56.9%; Score 2241; DB 13; Length 3813;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

QY      123  GGAAGTGAAGGGGAGGAAGGATGAGGAGCGCGCGCGCTGTGTGGCCGGCTCCACAC 182
Db      123  |||||
QY      194  GGGCGGGGCGGGGGGCTTGAGATGAGGCTTACCGCGGGGCTGTGTGCGCGCTCGATTA 253
Db      194  |||||
QY      183  CGCAACAGCTCGTGTGATCCGCGCGAGCGCGA-TCCCGGGCGGAAGCGCGCGGGA 241
Db      183  |||||
QY      254  CGGAACAGAGCTGTGTGTGATCCCGCGGACCGCGAGTCCGGAAGCCGCGGGCGCGCG 313
Db      254  |||||
QY      242  GCAGAACGGGCG---AGTGTGCCAATTTTCCGGCGACGAGTGTGGCTTCCCGCGGC 296
Db      242  |||||
QY      314  GCGCGCGCGGGCGAGAGCGCGGTGCAATATGCGGCGACGAGGTCGGGGTGGCTTGCAC 373
Db      314  |||||
QY      297  GGGGAGCCCTTGTGTGGGTGCAAGATGTGGCCCTCCCGGTCTCCCGGAGCTTGTAAGAA 356
Db      297  |||||
QY      374  GGGAGCCCTTGTGTGGGTGCAAGATGTGGCCCTCCCGGTCTCCCGGCTGTGACGAG 433
Db      374  |||||
QY      357  TACGAGCGCGCGGAGGCGACGACAGAACTGCCCGCAGTGCAGACTGCATACAGAGCCTC 416
Db      357  |||||
QY      434  TAGAGAGCGCGCGAGGCGCTCGCAAGCGTGCCTCGCGAGTGCAGAGACCCGCTAACAGCCTC 493
Db      434  |||||
QY      417  AAGGCGTCCCAAGCTGTGACCGGTGATGAGAGAGAGACGCGCGTGTGATGACTTGAGAACAC 476
Db      417  |||||
QY      494  AAGGCGTCCCGCGGAGTGGCCGAGAGAGAGAGAGAGCGCGTGTGACGACTTGAGAGGCGC 553
Db      494  |||||
QY      477  GAGTTTCAACTGG-----GACGGCGATGACTGCGAGTGTGTGGCGGAGTCC 521
Db      477  |||||
QY      554  GAGTTTGGCTGTGACGAGACGAGCGCGGCCACGAGAGCGAGCTCGAGTACTGCGGAGTCC 613
Db      554  |||||
QY      522  ATGCTTACGCGCCATGAGCTACGAGCGCGGTGAGGTGACCTTAAATGGCGCGCCACAAAGCT 581
Db      522  |||||
QY      614  ATGCTTACGAGGCGAGATGAGCTACGAGCGCGCGCGCGA-----CGGCGACCCCGGC 664
Db      614  |||||
QY      582  TTCCAGTCAACCCCAATGTTTCCACTCTCTACCCAAAGGCGCAATGTGTGATGACATCCCA 641
Db      582  |||||
QY      665  TTACGCGCGCGTCCCAAGTGCCTCTCTCAACAAAGCGCGAGATGTGATGACATCCCG 724
Db      665  |||||
QY      642  CCGGAGCAGACGCGGCTGTGGCTTCTTCTTCAATG-----GGTGTGGGGGAGAAAGAG 692
Db      642  |||||
QY      725  CCGGAGCAGACGCGGCTGTGGCTCTCTCAATGACGCGCGCGCGCGCGGGCGAGAGGG 784
Db      725  |||||
QY      693  ATACATCCCTTCTCTTATGCGGATCCAGCTTACCTGTGCAACCAGGCTATAGAACCCA 752
Db      693  |||||
QY      785  ATCAACCCGCTCTCTTTCGAGATCCCAACCTTCAAGTGCAACCGAGTTCATGAGACCG 844
Db      785  |||||
QY      753  TTCGAAGATCTTGTGTGATATGGGTATGTGTGTCTTGGAGAGAAAGCATGAGAT 812
Db      753  |||||
QY      845  TTCGAAGATCTGTGGCGCTTACGATATGAGCGCTGTGCGCTGGAAAGAGATGAGAGGC 904
Db      845  |||||
QY      813  TGAAGAGAGAGACAAAGAGAGATGACCAAGCGGAGATGATGTGGTGTGTGAT---GAT 869
Db      813  |||||
QY      905  TGAAGAGAGAGACGAGAGCGCTGACAGATGTACAGAGCAGAGGGTGGCGGTGATTTGGAT 964
Db      905  |||||
QY      870  GGTGACGATGTGTATCTTACCACTATAGATGAGCAAGACAAACAACCTGTCCAGAAAATT 929
Db      870  |||||
QY      965  GGCACAGATGAGATCTGCACTATAGATGAACACTAGGCGAGCAATGTCTCAGAAAAGTCT 1024
Db      965  |||||
QY      930  CCACTTCAATCAACCGAATTAATTCATATGATGATTAATTAATTTCCGCTGTGTGTT 989
Db      930  |||||
QY      1025  CCTATATCATACAGCCGAAATTAATCCCTACAGAGATGATATGTATTCCTGGTGTGTGTT 1084
Db      1025  |||||

```

OY	990	TTGGGGTTCCTCTTCCACTACCGAGTGTGATCCGGGAAATGATNGCATTTGGTTGGG	1069
Dp	1085	TTGGGTTTCCTCTTCCACTACCGAGTGTGATCCGGGAAAGATGCAATTTGCATTTGGG	1144
OY	1050	CTCATATCTGTATCTGTGAATCTGGTTTGCCATGTCTTGGAATCTTGATCAATGCCA	1109
Dp	1145	CTCATATCTGTATCTGTGAATCTGGTTTGCGATGTCTTGGAATCTTGATCAATGCCA	1204
OY	1110	AAGTGTTCCTTATTTGAGAGAGACTTTACCTAGACCGGCTGTCACTGAGGTTGCACAG	1169
Dp	1205	AAGTGTTCCTTATTTGAGAGAGACTTTACCTAGACCGTGTGTGTCACTGAAGTTGCACAG	1264
OY	1170	GAAGGCGCCCATCTCCAACTTGCTTCCTTGTGCACTAGCGGTGATCCCTTA	1229
Dp	1265	GAAGGTCACACCTCTCAGCTTGTCTCCAACTGACCTTGTGCACTAGCGGTGATCCCA	1324
OY	1230	AAGGAACCTCTTGTGTGACAAACAATCTGTTCATCTATCCCTTGCGTGGATTTATCCT	1289
Dp	1335	AAGGAACCTCTTGTGTGACAGCGACATCTCTTTCATCTCTTCTGTGGATTTATCCG	1384
OY	1290	GTTGTAAAGTTTCTTGTGATGTTTCTGTATGTGTGATGTGTGATGCTTAACSTTTGAACA	1349
Dp	1385	GTTGTAAAGTTCTCCTGTATGTTTCTGTATGTGTGATGTGTGATGCTTAACSTTTGAACA	1444
OY	1350	TTATCTGAAACATCTGAATTTTGCAAGAATGGGTTCTTTCTGCAACGGTACATATTT	1409
Dp	1445	TTGTCTGAAACATCTGAATTTTGCAAGAAATGGGTTCTTTCTCACAACAAAGTTTAAATC	1504
OY	1410	GAACCTCGCGCTCCAGAGTGGTACTTCCAAACAGAAATGACCTACTTGAAGAACAAGTG	1469
Dp	1505	GAGCTTCGTCTCTGATGTGATCTTCCAAACAGAAATGACCTACTTGAAGAACAAGTT	1564
OY	1470	GCAGCAAACTTTGTAGGAGAGAGAGACCAATGGAAGAGATGTGAAGATTTCAAGTG	1529
Dp	1565	GCTGCTTCATTTGTAGGAGAGAGAGGCGCATGGAAGAGAAATCGAAGATTTCAAGTTA	1624
OY	1530	AGATCAATGCTTATGTTGCCAAAGCCCAAGAAAGTTCTTGAAGAAGTGGACATGCA	1589
Dp	1625	AGGATCAATGCTTGTGTGCAAAAGCCCAAAAGGTTCTTGAAGAAGTGGACATGCAAA	1684
OY	1590	GATGGAAACCCCTGGCCTGGAAACAATGTTGTATCATCTCTGGAAGATTCAGGTCCTC	1649
Dp	1685	GATGGAAACCCCTGGCCTGGAAACAACATGACATCATCTGGAAGATTCAGGTTATTC	1744
OY	1650	CTTGGGCCAAACGAGAGCCTTGACCTGTGAGGGAATGACATGCAACATTTGGTTATGTT	1709
Dp	1745	CTTGGGCCAAATGGCGGTCGTGATGTGGAAGGAATGAGTTGCCCTCGCTGTATATGTC	1804
OY	1710	TCTAGAGAAACGACAGAGCTATTAACCATATAAGAAAGCTGTGTCTATGAATGACTTG	1769
Dp	1805	TCGAGAGAAAGAGGCGAGGTTATTAACCATACAAGAAAGCTGTGTGCATGAATGACCTG	1864
OY	1770	GTCGAGTCTCTGCTGTACTTAACAATGCTCATATTTGTTAACTTGGATTTGATGATC	1829
Dp	1865	GTCGGTGCTCTGCTGCTTAACAATGCTCATATCAATTAAGAACTTGGATCTGATGATC	1924
OY	1830	TACATCAACAACAGCAAGGCTATTAAGGAACAATGTGTTTATGATGGACCTTTACTA	1889
Dp	1925	TACATCAACAATGCAAGGCGCATTAAGAGGCTATGTGTTCAATGATGGATCTCTTGCTG	1984
OY	1890	GGAAGAAGGTTTGCTATGTACAGTTCCTCCAAAGATTTGATGGGATTTGATCGCATGAC	1949
Dp	1985	GGAAGAAGAAAGTGTGTATGTACAGTTCCTCCAGAGGTTTATGATGATTTGAACAATAAGAT	2044
OY	1950	CGATATGTGAACCGGAATGTGTCTTTTGTGATATCAACATGAAGGTTTGGATGTAT	2009
Dp	2045	CGATATGTGAACAGGAAGTGTCTTTTGTGATATCAACATGAAGGTTTGGAGGATTT	2104
OY	2010	CAGGATCAATTTATGTTGATGATGATGTGATTTGAAGGCAAGGCTATTAATGTTAT	2069
Dp	2105	CAGGATCCATTTATGTGGATCTGGATGTGTGTTTCAACGCGACGGCATCTTATAGTTAT	2164
OY	2070	GATGCCCCCAAAACAAGAACCAACATCAAGACTTGCACTGTGGCCCCAGTGTGTG	2129

Db	2165	GATCTCCTTAATAAGAAAGACCACATCAAGAACTTGCAACTGTGGCCCAAGTGTGC	2222
Oy	2130	TTTTCGTGTCGTCCTTTGGCAATGAGCAAAAGAAAGAACTACCAAAACCAAG	2183
Db	2225	CTCTCTTGCTGCTGCAGAGAGAAACAAGATTAAAAAGAAAGACTACAAAACCAAGCGAG	2288
Oy	2190	AAGAAAAAGTTATTATTTTTCAGAAAGAAAGAAACCAATCCCTGCATATGCTCTTGST	2245
Db	2285	AAGAGAAAGAAAGATTATTTTTCAGAAAGAAAGAAACCATCTCTGCATATGCTTGSGST	2344
Oy	2250	GAAATTGACGAGCTGCTCCAGAGCTGAGAAATGAAAAGGCGGTAATTGTAAATCAACA	2304
Db	2345	GAAATTGATGAGGTGTCTCCAGGCTGTAATTCAGAAAGGCGGTAATGTAAATCAAG	2404
Oy	2310	AAATTAGAAAAGAAATTTGGCCAACTCTTGTTTTGTATCATCCACACTCTCGAAGAT	2366
Db	2405	AAACTGAGAAAGAAATTTGGCCACTCTTGTTTTGTGCGATCAACACTCTTGAGAAC	2466
Oy	2370	GGTGGAACCTTGAAGAGTCAGAGTCTGCTTCTCTTTGAAAAGAGCTATACATGTCATT	2422
Db	2465	GGAAGGACCCCTGAAGAGGCAAGTCCAGCTTCTCTTGAGAGAAAGCTATACATGTTATC	2522
Oy	2430	AGTTGTGTTATGAAGACAGACAGACTGGGGAAGAAAGATTTGGCTGGATCTATGATCA	2488
Db	2525	AGCTGCGCTACGAAACAAAGACGACTGGGGAAGAAAGATTTGCTGGAATTTACGATCG	2584
Oy	2490	GTTACAGAAAGATATTCTTAACGTGTTTCAAGATGACTGTCAATGGTGGCGGTCAATTAC	2546
Db	2585	ATCACAAGAGATATCTTGACTGGAATTTAAGATGACACACACAGCTGGCGGTCTATTTAC	2644
Oy	2550	TGCATACCTAAACGGGTGTGCAATTCAGAAAGTTCGCACTCTGAAATCTTTTCAGATGTCCT	2606
Db	2645	TGCATCCCGAAGCGGCTCGCATTCAGAAAGTTCGCGCTCTGAACTTTTCCGACGCTCTT	2704
Oy	2610	CACCAAGTGTTCGGGTGGGCTCTTGGGTCTATGAGATCTTCTTCAGAAATCATTTGCCCT	2666
Db	2705	CACCAAGTCTTTCGCTGGGCCCTTGGGTCGCTGAAATTTTCTTCAGCAACACATGCCCA	2766
Oy	2670	CTTTGGATGGGTATNGGTGGCGGTCTGAAATTTTGTGAAAAGATTTTCTCATCAACTCC	2722
Db	2765	CTTTGGATGGGATACGGCGCGCGGCTAAATTTCTGGAAGGTTTTCTTATATCAACTCC	2822
Oy	2730	ATCGTGTATCTTGGAGACATCTAATTCCTCTTGCTTACTGTACATGGCTCCGCATCTGT	2788
Db	2825	ATCGTGTATCTTGGAGAGTCAATTCCTCTCTGCTTACTGTACATGGCTCCGCATCTGC	2888
Oy	2790	TTATTTGACAGGAAATTTATCATCTCCAGAGCTGAATATGTTGCCAGCTGTGTCTATG	2844
Db	2885	CTGCTCAACGGGAAAGTTTATCACACACAGAGCTTACCAATGTCCGACATCTGTTTCAATG	2944
Oy	2850	TCACCTTTTATCTGCAATTTTGTGTAACAGACATCTTAAGAAATGAAATGAAATGAGTGTGGA	2904
Db	2945	GCACCTTTTCACTGTCACTCCGTGACCGGACCTCTGGAATGAGGTGAGTGGCGTGTCC	3004
Oy	2910	ATTGATGACTGTGAGAGATGAGACAGTCTTGAGTCAATTGAGGNGTGTCTCAACACTC	2966
Db	3005	ATGACACACTGTGTGAGAGAACGACAGTCTTGAGTCAATGAGGCTTTCCGCGCATCTG	3066
Oy	2970	TTTGTGTGTTCCAGGGACTTCTCAAGGTCAATGCTGTGTGTTATACAAAGCTTCAACGTTG	3022
Db	3065	TTCCGCGGTGTTCCAGGCGCTGTGAAAGTGTTCGCCGCGATCGACACAGACTTCAACGTTG	3122
Oy	3030	ACATCAAAAGGTGAGATGATGAGAGAGTCTCAAGCTATATACATTCMAATGACATAC	3088
Db	3125	ACGTGAAAGGCGGAGAGCAAGAGAGTCTCGAGAGCTATACAGTTCAAGTGAACAC	3184
Oy	3090	TTATTTGATACCTCTCAACACTTGCTTATTTGAACTTCATTTGGTGTGTGCTGTGGCTT	3146
Db	3185	CTGCTGATATCCCGGACCAACGCTCTCTGCTGTGAATCTTCAACGGGAGTGTGCGGAGATC	3244
Oy	3150	TCAATGCGATCAATTAACGATATGAGTCATGTGGGCCCTCTTTTGGAAAGCTATTTCTTT	3209

Query Match	Match Similarity	38.3%	Score 1509;	DB 10;	Length 3255;
Best Local	Similarity	67.6%	Pred. No. 0;		
Matches 2207;	Conservative	0;	Mismatches 1030;	Indels 30;	Gaps 5
Qy	144	ATGAGAGCGAGCGCCGGGCTGGTGGCCGGGCTCCCAACACCGAACAGCTCGTCGATC	203		
Db	1	ATGAAACACCGGTGGTGGTTAATCGCCGGTTCCACCAACAGAAATGATTTGGCTCATTT	60		
Qy	204	CGCCGCGAGCGCGATCCCGGCGCGAAGCCGCGGGAGACAGAACGGGCAAGTGTGCCAG	263		
Db	61	AATGCCGATGAAGAAATGCCCGAATTAAGATCAGTCTCAAGAACTGATGGACAGACATGTCAA	120		
Qy	264	ATTTCGGGAGAGACGTGGGCTTGGCCCCGGGGGAGACCCCTGTGGCGTGCACAGAG	323		
Db	121	ATCTCAGAGATGAGATGCAATTTGATGTTTATGAGAAACCTTTGTGTGCAATGTAACGAA	180		
Qy	324	TGCGGCTTCCCGTCTGCGCGGAGCTGTACGAATACAGACGCGGGAGGGCAGCGAAGC	383		
Db	181	TGTGATTCCTCTGTGTGTAGACCTTGTCTATGAGTACGAAGAAGAGAAAGCAATCAAGCT	240		
Qy	384	TGCCCCAGTGCAGAGCTCGATACAAAGCCCTCAAGGGCTGCCAAGTGTGACCGGTGAC	443		
Db	241	TGTCCACAGTGCAGAAACCGTTTCAAAACCTTTAAAGAGATCCAAAGATTGAGAGGTGAT	300		
Qy	444	GAGAGAGAGACGGCGTCATGACCTCTGGACCAACAGATTTCACTGGGACCGGCATGACTCG	503		
Db	301	GAAAGGAAAGATGACATTATGATTTTAAACAAATGAGATTTTGA--GTATGAAATTAATGGG	357		

OY	504	CAGCTGTGGCCGAGTCCATGCTCTTACGGCCACATGAGCTACCGGCCGTGAGGTGACCT	563
Db	358	ATTGGAATTGATCAGTTTCTGAAAGGATATGTCATCTCTGTCGCAACTCCGGTTTCCCA	417
OY	564	AATGGCGCGCCACAGCTTTCACGCTCAACCCCAATGTTCCACTCTCACCACGGGCAA	623
Db	418	CAATCTGATTTGGATTGAGTTCAGCTCCACCTGGGCTCTAGATTTCCATTGCTGACTTACGGGAC	477
OY	624	ATGGTGATGACATCCACCGGAGCAGCAGCGC---TGGTGCCTTTTCATGGGTGT	680
Db	478	GAGAGACGTTGGAGATTTCTTCGTATGACATGCTCTTATTTGTTCCCTTCCTTCAGTGTGT	537
OY	681	GGGGGAAAAGAGATACATCCCTTCCTTATGCGGATCCGAGCTTACCTGTCGCAACCCAG	740
Db	538	CATGGCAATAGAGTTCACTCCTGTTTCTCTTTCTGACCCGACCGGTGCTGACATCTCAAG	597
OY	741	TCATATGACCCATCCAGATCTTGTCTGCAATATGGATATGGTATGTTGCTTGGAAAGAA	800
Db	598	CCATATGTAACCTCAGAAAGATCTTCGGGTTTATGGTATGGAAGTGTGCTTGGAAAGAT	657
OY	801	CGGATGAGATTTGGAAAGCAGAGACAAAGAGATGACCCAGCGGGG-----AT	851
Db	658	CGAATGAGGATGGAAAGAGAAAGACAGATGAGAACTTCAGTTGTTTGGCATGAAAGGA	717
OY	852	GATGGTGTGTGATGATGTGTGACGATGCTGATCTACCACTTAATGATGAAAGCAAGCAA	911
Db	718	GATCCTGATTTTGAAGATGATGATGATGATGATTTTCCAAATGATGATGAGGAAGGAG	777
OY	912	CAACTGTCCAGAAAATTTCCACTTCCTCATCAAGCCAGATTAATCCATATAGATGATATC	971
Db	778	CCATTGCTTAGGAAAGATACCAATCAATTCAGCAGAAATTAATCTTACCGGATGTATTT	837
OY	972	ATTATTCGGCTTGATGTTTGGGGTCTTCTTCCTCAACGAGATGATCCGGTGAT	1031
Db	838	GTGCTAGCTCTTGATGATTTCTGTCTCTTTCTTCACTACGATATTTCCACCCGTCAA	897
OY	1032	GATGCAATTGCTTTGTGGCTCATATCTGTATCTGTGAATCTGTGTTGCCATGTCTTGG	1091
Db	898	GATGCAATATGCTTTGTGGCTTATTTCTGTATATGTGATATGATGATGTTGCTGTATTACGG	957
OY	1092	ATTCTGTATCAATTTCCCAAATGCTTCCCTTATGAGAGAGACTTACCTTACACCGGCTG	1151
Db	958	GTCTCTTATGACGTTCTCCCTTAATATGATACCTTATCGACGAGAAAGTATTTGACCCGACTC	1017
OY	1152	TCACTGAGGTTGCAACAGGAAGGCGAGCCATCTCAACTGTGCTCAATTTGATTTCTTGTGC	1211
Db	1018	TCATTAAGATATGAGAAAGAAAGGGAACCGTCGGGACATATCCCGTGGAATGTAATTTGTT	1077
OY	1212	AGTACGCTTGAATCCCTTAAAGAACTCTCTTTGGTACACAAATACTCTGTCTATCTATC	1271
Db	1078	AGTACAGGATATCAATTGAAAGAGCTCCCGCTTATATCTGCAAAATACGTCTTGTCTAAT	1137
OY	1272	CTTTGCGTGAATTAATCCTGTGATTAAGGATTTCTTGCTATGTTTTCGATGATGCTGCTCA	1331
Db	1138	CTTGCTGTGATATCTCTGTGATTAAGGCTCTTCTTACGATCTGATGATGATGCTGCTCT	1197
OY	1332	ATGCTAACGTTTGAAGCATTAATCTGAAACATCTGAATTTGCAAAAGAAATGGGTTCTTTC	1391
Db	1198	ATGCTTATCTTTCGAAAGCTCTTCTGAGACCGCTGAAATCGCAAGAAATGGGTTCTTTC	1257
OY	1392	TGCAAAACGTTCAATATGAAACCTCGGCTCCAGAGTGTACTTCCACAAAGATTAAC	1451
Db	1258	TGCAAGAAATATTTGATTAAGGCTCTGCTGCCGATGATATTTCTGCAATAAATGAGAC	1317
OY	1452	TACTTGAAGAACAAGGTGGCAGCAAACTTTGTATGGGAGAGAGAGACATGAGAGAG	1511
Db	1318	TACTTGAAGATTAAGTTCATTCGCCGATTTGTATAGGAGCGCGGAGCATGAGAGAGAT	1377
OY	1512	TATGAGGAATTCAGGTGAGATCAATGCCCTTAATGTGCAAAAGCCCAAGAAATGTTCTGAA	1571
Db	1378	TATGAAGAATTCAAAGTAAAGATCAATGCTTTAGTAGAACACAGCAACAAAGTGTCTGAG	1437

QY	1572	GAAGGATGAGCAATGACAAGATGAAACCCCTGGGCTGGAAACAAATGTTGCGATCATCTT	1631
Db	1438	GATGGTTGACCTATGACAGACGGTACACTTTGGCCGGTAATAGTGTGCGAGATCATCT	1497
QY	1632	GGAAATGATTCAGGCTCTTCTTGCGCAAAAGCGAGGCGCTTGACTGTGAGGGAAATGAAC	1693
Db	1498	GGCATGATTCAGGCTCTTCTTGGAAGTACGGGTGTTGTTGATGTGCAAAACACAGAGTGG	1557
QY	1692	CCACGATTTGGTTTATGTTTCTAGAGAGAAACGACCAGGCTTTAACCATCATTAAGAAAGCT	1751
Db	1558	CTTCGATTAGTTTACGTTTCTCGTAGAAGAGACCCGGAATTTGATTCACCATTAAGAGGCT	1617
QY	1752	GGTGCCTATGAAATGCAATGTGTCGGAGTCTGTGCTGTACATAAAGATGCTCCCATATTGTTA	1811
Db	1618	GGAGCTATGAAATTCCTGTATACGAGTCTGTGGGTTCTATCAAAATGCTCTTACCTTCTG	1677
QY	1812	AACCTTGATTTGATTCACCTACATCAACACAGCAAGGCTATTAAGAAAGCATGTGTTT	1871
Db	1678	AATGTGCATTTGTATGATCACTACATCAACATAGCAAAAGCTCTTAGAAGACATGTGTTTC	1737
QY	1872	ATGATGGAACCTTTTACTAGAAAGAGGTTTGTATGTACAGTTCCCTCAAAAGATTGAT	1931
Db	1738	ATGATGATCTCTCAGTCAGAAAGAAATCTGTATGTTCAGTTCCCTCAAAAGTTGCAT	1797
QY	1932	GGGATTTGATGCGCCATGACCCATATGCTAACCGGAATGTGTCTTTTGTGATTCACATG	1991
Db	1798	GGGATTTGATGACCATGATCGATCTCAATGTGCATATGTGTGTTCTTGTGAATCATATGT	1857
QY	1992	AAAGTTTGAATGAGTAATCAGGGTCCAAATTTATGTGTACTCGGATGTGTATTAGAGG	2051
Db	1858	AAAGTTTGAATGAGGCTACAGGGCTATATACGTGGTACAGTTGTGTTCAGAGG	1917
QY	2052	CAGGCATTATATGTTATGATGCCCCCAAAACAAAGAACCATCAAGACCTTGCAAC	2111
Db	1918	CAACGCTTTACGATTTGATGCAACCGAAGAAAGAGAGGCGCCACGTAAACATGCAT	1977
QY	2112	TGCTGGCCCCAAGTGTGCTTTTGCTGTGTCTGCTTTGGCAATAGGAACAAAGAGAGCT	2171
Db	1978	TGCTGGCCAAATAGGTGTCTCTATGTGTTTGTTCCAAAGAAATCTGTAAAGCAAGCA	2037
QY	2172	ACCAAAACCAAAACAGAGAAAGAAAGTATATTTTCAAGAAAGAGAACCAACAC	2231
Db	2038	GTGGCTGGGATTAAGAAAGAAAGAA-----TAGGAAAGCTCAAGCAATC	2085
QY	2232	CCTGCATATGCTCTTGTGTAATTTAGCAAGACCTGCTCAGAGAGCTGAGATGAAAGGCC	2291
Db	2086	CAGCATTAGAAAATATGGAAGAGGCGCGTCACTAAGTTCTTAAAGTATGAACAGTCA	2145
QY	2292	GGTATTTGAATCAACAAAATTAAGAAAGAAATTTGGCCAACTCTTGTTTTGTTACA	2351
Db	2146	ACCGAGGCAATGCAAAATGATGGAGAAATTTGGCAGTGTCTGTATTTGTGCA	2205
QY	2352	TTCACACTTCTCGAAGATGTGTGAACCTTGAAGAGTGCATGCTCTTCTTTTGAA	2411
Db	2206	TCTCCGGCTATGGAAGAAAGTGTGGAGTGTCTGAAACGCAACCCGCGTGTCTGTGTTAA	2265
QY	2412	GAACTATATACATGTCATTAAGTTGTGTATGAAAGCAAGACAGACTGGGGAAAGAGAT	2471
Db	2266	GAAACCATCAAGTCATTAAGTTGGGATATGAAGATTAACCTGAATGGGGAAAGAGAT	2325
QY	2472	GGCTGGAATCTATGGAATCAGTTACAGAGATATTTCTAACCTGTGTTCAAGATGCATGTGCAT	2531
Db	2326	GGGTGGAATCTATGTTGTGTATACGGAATATTTCTTAAGGGTTTTAATGATCATTTCTCAT	2385
QY	2532	GGTTGGCGGTCAATTTTACTGCATACCTTAAAGGGTTCATTCMAAGTTTCTGCACCTGTG	2591
Db	2386	GGTTGGAGATCTGTTTATTTGTATACCAAAAGTTAAGGGCTTTCAAGAGATCAGCTCCATTC	2445
QY	2592	AATCTTTAGATTCGCTTACACAGGATGCTTGGGTGGGCTCTTGGGTCTAATGAGATCTTC	2651
Db	2446	AATCTTTGGAATCGTCTCATCAACAGTTCTTGTGATGGGCGCTTGGGTGCGTTGAGATTTTC	2505
QY	2652	TTACGCAATCATGGCCCTCTTGTGTATGGGATATGCTGGCGGTCTGAAATTTTGTGGAAGA	2711

Db	2506	TTGAGTACGAGATGTGTCTTATTTGGTATGGTATGAGAGTGGGTGAAAATGGCTTGACCGG	2565
Qy	2712	TTTTCTCAATCACTCACTCCATCGTGATCCTTGAGACATTAATCCCTCTTGCTTAACGT	2771
Db	2566	TTGTCTCACTAATCACTGTGGTTTACCCGTGAGCCTCTCAACCGCTCAATCGTTTACTGT	2625
Qy	2772	ACATTCGCTCCGACATCTGTTTATATGACAGAGAAAATTATACATCCAGACGTGAATTAATGT	2831
Db	2626	TTCTCTCCCTGCACATCTGTCTTCTCACTGGAATAATATGATGTTCGCAAGATTAAGCACTAT	2685
Qy	2832	GCCAGCCTGAGTTCATGTCATCTTTTATTCGTGATATTTTGTCATGAGACATCCAGAAATG	2891
Db	2686	GCGAGTATCTCTTCAATGGCGCTCTTCTGTGATTCATTAACGGGTATTTCTCGAGATG	2745
Qy	2892	ACATGAGTGGTGTGAAATTGATGCTGATGAGAGATGAGACAGTTCTGGGTCAATTGGA	2951
Db	2746	CAATGGGGCAAAAGTTGGGATCGATGATGATGGTGGAGAAAGAACAGTTTGGGTCAATTGGA	2805
Qy	2952	GGGTGTCTCTCAACCTCTTGTGTGTTCAGAGGACCTTCAAGGTCAATAGCTGTGTGT	3011
Db	2806	GGGTGTCTCTCGCATGTGTGTGTCTCTTCCAAAGTCTCTCAAGGTCTTGTGTGTGTGT	2865
Qy	3012	GATACAGCTTCAACCGGACATCAAGAGGAGATGATGAGGAGTTCACAGCTAATAT	3071
Db	2866	GACACTTAATTCACAGTCATCATCAAAAGCAGCTGATGATGAGAGTTCCTGTGACCTTTAC	2925
Qy	3072	ACATTCAAATGACATCACTTATTTGATACCTCTCAACGCTTGCTTCAATGAACTTCAAT	3131
Db	2926	CTCTTCAAAATGACATTCATCTTCAATCCCTCGAATGATCTCAATCATTAACGTCAAT	2985
Qy	3132	GGGTGTGTCTGCGGCGCTTTCAAATGCGATCAATTAACGATATGATGATGAGGCGCCCTC	3191
Db	2986	GGAGTCAATAGTCGAGATCTCTGTATGCCATACAGAAATGATTAACAGCTGTGGGAGCCGCTT	3045
Qy	3192	TTTGGGAGCTATTTCTTTGATTTTGGGTATGTTCATCTTATTCCTTTCTCAAAAGT	3251
Db	3046	TTGGGAAAGATGTCTTTTGGACTTGGGTATCATTAATCAATCCGTTCTTAAAGGT	3105
Qy	3252	TTGGTGGAAAGCAAAACAGACACCAAGATGTGATGCTGTGTCATTTGCTGGCT	3311
Db	3106	TTGCTTGGGAAACAAAGATGAATGAATGCCAACATTAATGTGCTGTGTCATCTCTCGGC	3165
Qy	3312	TCAATCTTCTGCTCTCTTGGGTGGGTGAGATTTGATCTTTCTTGGGAGAGATGATGTCCG	3371
Db	3166	TCGATCTTCACTTCTTGGGTCCGGGTAAATCCGTTGTGGCGA---AAGCGGTCTC	3222
Qy	3372	CTTCTTGAGAGATGTGTGTTGGAATGC	3398
Db	3223	ATTCTCGAAGATCTGTGTTTGAAGCTGC	3249



SEQ ID NO 559  
 LENGTH: 3255  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-559

Query Match 36.4%; Score 1431.8; DB 10; Length 3255;  
 Best Local Similarity 66.0%; Pred. No. 0;  
 Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

QY 144 ATGAGGCGAGCGCGGCTGTGGCCGCTCCCAACCGGAGAGCTGCTCATC 203  
 DB 1 ATGAATACGTGTGTGCGCTCATTTGCTGCTCTCAACAGAAAGAAATGTTCTCAT 60  
 QY 204 CGCCCGAGCGCGATCCCGGCGGAAGCCGCGGAGCAGAACGGGAGGTGCCAG 263  
 DB 61 AACGCGATGAGAGTGCCAGAAATACGATGACAAAGAACTGAGGGGCAACATGTCAA 120  
 QY 264 ATTTCGCGGAGAGAGTGGGCTTGTGCCCCGGGGAGCCCTTGTGGCGTCAACGAG 323  
 DB 121 ACTGTGAGATGAAATCAATTTAACGTTAGCAGTGAAGCTTTGTTGCTTGCAACGAA 180  
 QY 324 TCGGCTTCCCGCTGCGGGAGCTGTAGATACGAGAGCGCGGAGGGGACGAGAAC 383  
 DB 181 TCGGCTTCCCGCTTGTAGACCATGTCTATGATATGACATGAGAGAGAAATCAAGT 240  
 QY 384 TCCCGCCAGTGCAGACTGTGATACAGCCGCTCAAGGGCTGCCAAGTGTGACCGGTGAC 443  
 DB 241 TGTCTCAGTGCAGAACTCGATACAAAGATTTAAAGGTAGTCCACGGGTTGATGGAGAT 300  
 QY 444 GAGGAGAGAGAGCGGCTGATGACCTGGACACAGATTCACTGGAGCGGCATGATCTG 503  
 DB 301 GATGAGAGAGAGAGAGACATTTGATGATCTTGAATGATGATTTGATCATGGAGTGAACCT 360  
 QY 504 CAGTCTGTGCGGAGCTGATGCTTACGCGCACATGAGCTTACGCGCGTGGAGTGAACCT 563  
 DB 361 GAAATGCGGCTGAGAGAGCGGACCTCTTCAACGCTTACACCGGCTGTGGTGAATTTGAT 420  
 QY 564 AATGGCGCGCAAGAGCTTTCAGCTCAACCCCAATGTTTCACTCTCAACCAAGCGGCA 623  
 DB 421 TCAGCTCCACCTG-----GCTCTCAGATTCCTCTTTGACTTATTGTGAT 465  
 QY 624 ATGTGTGATGATATCCACCGGAGCAGCAGCGCTGTGCTTCTTCAATGGGTGTGG 683  
 DB 466 GAAATGCTGATATGATTTCTGATCGTCAATGCTTATGTGCTCTTCAACGGGATAT 525  
 QY 684 GGAAGAGGATATCATCCCTTCTTATGCGGATCCCAAGCTTACCTGTGACACCGAGTCT 743  
 DB 526 GGAATGCGGCTATCTCTGCAACGTTTACAGATTTCTTGCACCTTCAACGGCGAGATCA 585  
 QY 744 ATGACCCATCCAGAGATCTTGTGATATGGGTATGTAGTGTGCTTGAAGGAAACG 803  
 DB 586 ATGCTTCTCAGAAAGATATTTGGGAAATATGGTATGGAAGTGTGCTTGAAGGACG 645  
 QY 804 ATGAGAAATTTGAGAGAGAGACAGAGAGATGACACAGCGGAAATATGTGTGT 863  
 DB 646 ATGGAAGTTTGAAGAGAGACAGCAAGGCAAAAGCTTCAAGTCAATTAAGCATGAAGAGGA 705  
 QY 864 GATGATGT-----GACGATGCTGATCTACCACTAATG 896  
 DB 706 AACATGTGTGAGGTTCCATGATGACGAGAACTAGATGATCTTACATGCTATGATG 765  
 QY 897 GATGAAGCAAGACAACTGTCCAGAGAAATTTCCACTTCAATCAAGCCAGATTAATCA 956  
 DB 766 GATGAAGGAAGAACCTCTCTCAAGAAAGCTACATTTGTTCAAGAGAAATTAATCT 825  
 QY 957 TATAGATGATATATATATGCGCTGTGGTTTGGGTTCTTCTTCACTACAGAGG 1016  
 DB 826 TACAGAGATGATATTTCTGTGCTGCGGATTTCTGTGCTTTTCTTATTAATTAAT 885  
 QY 1017 ATGATCCGCTGATGATGATGCTTTGTGGCTCATATCTGTATCTGTGAATCTGG 1076  
 DB 886 CTCATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 945

QY 1077 TTTGCGATGTCTTGATTTCTTATCATATCCCAAGTGTCTCCTATTGAGAGAGACT 1136  
 DB 946 TTTGCGATGTCTTGATTTCTTATCATATCCCAAGTGTCTCCTATTGAGAGAGACT 1005  
 QY 1137 TACCTAGACGGGCTGTCACTGAGGTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196  
 DB 1006 TACCTAGATGATCTCTCTCAAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065  
 QY 1197 ATTTATTTCTTTGTCAGTACGTTGATCCCTTAAAGAACCTCTTGTGTCAACAAT 1256  
 DB 1066 GTTATGATTTTGTATGACAGTGTGATCCGTGAAAGAGCAGCAGCTTGTATGAGCAAC 1125  
 QY 1257 ACTGTTATATATCTTTGCGTGTGATATCTGTGTATTAAGTTTCTGTATGTTCT 1316  
 DB 1126 ACAGTTCTTTCATCTAGCAGTTGATTAATCTGTGATTAAGTTGCGTGTATGATCA 1185  
 QY 1317 GATGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376  
 DB 1186 GACGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1245  
 QY 1377 AATGCGTTCTTTCTGCAACGTTACATATTTGAACCTCGCGCTCCAGAGTGTACTTC 1436  
 DB 1246 AATGCGTTCTTTCTTTGTAAGAAATTTATATGAGCAGAGAGCTCCTGAGTGTATTT 1305  
 QY 1437 CAACGAAATGATCTACTTGAAGACAGGTGACAGCAACTTTGTTAGGAGAGAGA 1496  
 DB 1306 TCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1365  
 QY 1497 GCAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556  
 DB 1366 GCTATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425  
 QY 1557 CAGAAAGTTCTTGAAGAGATGAGCAATGAGATGAGAACCCCTGAGCTGAGAACAT 1616  
 DB 1426 CAGAAAGTTCTTGAAGAGATGAGCAATGAGATGAGAACCCCTGAGCTGAGAACAT 1485  
 QY 1617 GTTGTGATCTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1676  
 DB 1486 GTTGTGATCTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1545  
 QY 1677 GAGGAAATGAACTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1736  
 DB 1546 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605  
 QY 1737 CATCATGAAGAGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1796  
 DB 1606 CACCAAGAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665  
 QY 1797 GCTCATATTTGTTAACTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1856  
 DB 1666 GCTCATATTTGTTAACTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1725  
 QY 1857 GAGCAATGTTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1916  
 DB 1726 GATTTATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785  
 QY 1917 CCTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1976  
 DB 1786 CCGCAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845  
 QY 1977 TTTGATATCAATGAAGGTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 2036  
 DB 1846 TTTGATATTAATGAGAGGTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1905  
 QY 2037 TGTGATTTTGAAGGAGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 2096  
 DB 1906 TGTGATTTTGAAGGAGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1965  
 QY 2097 TCAAGACTTGCACCTGTGCGCAAGTGTGCTTGTGCTGTGCTGTGCTGTGCTGTG 2156  
 DB 1966 GGCAGAAACCTGTAATGTTGCTTAATGATGATGATGATGATGATGATGATGATGAT 2025







1072 GTGGCTAGATCTGTTATCTGAGGCTGGTTGCTTGTCTGCTTCTAGATCAGTT 1131  
1106 CCCAAGTGTTCCTTATGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTGA 1165  
1132 CCCAAATGATTCACATCAACCGTGAAGACATATCTCGACAGGCTTGCAATGAGTAA 1191  
1166 CAGGAGGCCGACCATCTCAACTTGTCTCAATTTGATTTCTTTGTCAAGTACGGTTGATCC 1225  
1192 TAGAGAGGAGGAGCCATCAAGCTGCTCCCATGTATGTCCTTGTGACAGTACAGTGAATCC 1251  
1226 CTAAAGGAACTCTTTTGTGACAACTACCTGTCTATCTATCTCTTGGTGGATTA 1285  
1252 ATGGAAGAACTCCACTGATCAAGCCAACTGTGTTGTCTCAATCTTGTGCTGATTA 1311  
1286 TCCGTGTATAGGTTTCTTGTCTATGTTCTGTGATGTGTGTCATGTCATACCTTTGA 1345  
1312 CCGTGTGACAAAGTGATGATGTTATGTTCTGTATGATGCTCAGCTATGCTGAATTTGA 1371  
1346 AGCATTTATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACGGTACAA 1405  
1372 GTCTCTCTGTGAACCTGCGAATTTGCTGAAGAGTGGGTTCCCTTTGTGAAGACACAA 1431  
1406 TATGAACTCGCGCTCCAGATGGTACTTCAACAGAAATAGACTTGAAGAGCAA 1465  
1432 TATTAACCAAGAGCTCCAGAAATTTACTTGTCTCAAAAATAGATTACCTGAAGACAA 1491  
1466 GTGCGACGAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGTATGAAGAAATTCAA 1525  
1492 AATTCACCTCATTTGTTAAGAAAGAGACCAATGAAGAGATGAAGAAATTCAA 1551  
1526 GGTGAAGATCAATGCTTATGTTCCAAAGCCCAAGAAAGTTCTTGAAGAGATGACAT 1585  
1552 AATAAGAAATCAATGCTTGTGTGCAAGACAGAAAGTGCCTGAAGAGGCTGACAT 1611  
1586 GGAAGATGGAACCCCTGGCTGGAACCAATGTGTGATCATCTCGGAATGATTCAGAT 1645  
1612 GCTGATGGAATCTGCTGGCTGGAGATTAACCTTGAAGACATCTGGATATTTCAAGGT 1671  
1646 CTTCCTTGGCCAAAGCGAGGCTTGACTGTGAGGGAATGAATGCCACGATTTGATTA 1705  
1672 GTTCTGGGGGACAGTGGGGCTTGACACTGATGGAATGAATTAACACGCTTGTCTA 1731  
1706 TGTTCCTAGAGAAACGACAGGCTATACATCATTAAGAAAGCTGTGTATGAATGC 1765  
1732 TGTCTCTGTAAGAAAGACAGGCTTTCAGCATCAAGAAAGCTGTGTGCAATGAATGC 1791  
1766 ATTTGCTCGAGTCTGCTGTACTAACAATGCTCATATTTGTAAATTGTGATTTGA 1825  
1792 ACTGATTCGTATCTGCTGTCTGACAAATGCTGTATCTTCAATGTGATTTGA 1851  
1826 TCACATCATCAACAACAGCAAGGCTATAAAGAGCAATGTTTATGATGACCTTT 1885  
1852 CCATTACTTCAATAGAGCAAGCTCTTAGAGAGCAATGTCTTCATGATGATCCAGC 1911  
1886 ACTAGAAAGAAAGTTTGTCTATGTACATTTCCCTCAAGATTTGATGAGATGATGCCA 1945  
1912 TCTAGGAAGGAAACCTTGTATGTACAAATTTCCAAAGATTTGATGCAATGTGAC 1971  
1946 TGAACGATATGCTAACCCGAATGTTGTCTTTTGTATATCAACATGAAGGTTGAGATGG 2005  
1972 CAGTGAATATGCTAAAGGAACATGTCCTTTGATATCAACATGAAGGCTGATGATGG 2031  
2006 TATTCAGAGTCAATTTATGTTGTTGTTACTGATGTGTATTTAGAAAGCAGGCAATTAATGG 2065  
2032 CATTCAGAGTCCAGTCTATGTTGGAACAGAGATGCTGTTCAATPAGCAGGCTTTGTATGG 2091  
2066 TTATGATGCCCCCAAAACAAAGAGCCACATCAAGAGCTTGCAACTGTGCCCCAAGTG 2125  
2092 ATATGAT-----CCTGTTTGAAGTGAAGCTGATCTGGAACCTTAACAT 2134  
2126 GTTCTTTGCTGTGTGCTTTGGCAATAGGAAGCAAGAAAGATCAACCAACCAAAAC 2185

2135 GTTGTAA-----AGCTGTGTGTAGAGAGAAAGAAACAAAGATTATATGATAG 2190  
2186 AGAGAGAAAAAGTTATATTTTCAAGAAAGAGAACCAATCCCTTCATATGCTCT 2245  
2191 TCAAGCCGT-----TATGAGAGAAACAGATCTTCAGCTCCCATCTTTAATCAT 2241  
2246 TGTGAATTTGACCAACTGCTCCAGAGAGCTGAGAAATGAAGAACCCGCTATTTGAATCA 2305  
2242 GGAAGACATCGAGAGGATATGAAGTTATGAAGATGAAGAGTCAAGTGTATGTCCCA 2301  
2306 ACAAATAATGAAGAAATTTGSCCAATCTTCTGTTTGTATCATCCACATCTTCA 2365  
2302 GAGAAATTTGAGAAAGAGCTTTGCTAGCTCTCCATCTTCAATTCACATCCATTAAGAC 2361  
2366 GAATGTGGAACCTTGAAGAGTGAAGTCTGCTTCTCTTTTGAAGAGCTATACATGT 2425  
2362 TCAAGGTGACATCACCTTCAACAAACCCAGCTTCTCACTGAAGAGATATCAATGT 2421  
2426 CATTAATTTGTGTTATGAAGCAAGACAGACTGGGGAAGAAAGATTTGGCTGATCTATGG 2485  
2422 TATCAGCTGTGGTACAGAGCAAACTGAATGGGAAAGAAAGATTTGGCTGATCTATGG 2481  
2486 ATCAGTTACAGAAATATTTCTAATCTGTTTCAAGATGCAATTTGATGTTGGCGTCAAT 2545  
2482 TTCACTTACAGAGATATTTCTGACTGGGTTTAAATGCAATGCAAGAGCTGGCAATCAT 2541  
2546 TTACTGCATACCTAACCGGTTGCATTCAAAGTTCTGACCTCTGAATCTTTCAATCG 2605  
2542 CTACTGATGCAACCAAGACCTTGTTCAAAGGTTCTGACCAATCAATCTTTCTGATCG 2601  
2606 TCTTACACAGTGTGTTGGTGGGCTCTTGGGCTATGAGATCTTTCAGCAATCATG 2665  
2602 TCTTAATGAGTGTCTCGTTGGGCTCTTGGGCTGAGTGAATTTGCTTGAACAACATG 2661  
2666 CCTCTTTGTGATGAGTATGCTGCGGCTGAAATTTTGAAGATTTTCTCATCACTAA 2725  
2662 TCTTAATGATGATGCTGACATATGGGCAATGAACCTTTGAGAGGCTGCTTACATTA 2721  
2726 CTCATGCTGATCTCTTGAACATCTATCCCTCTTGGCTTACTGATCATGTCCTGCAT 2785  
2722 CACCATGTTTATCCAAATCAATCATCTGTTCCGCTTATGCGCTATGTGTGCTTCCGCTAT 2781  
2786 CTGTTTATGAACAGGAAATTTATCACTCCAGACTGAATTAATGTTCCAGCCGTGTT 2845  
2782 CTGCTCTTCAACAATAATTTATCATTTCTGAATTAATGAATATCTGGAATGTTCTT 2841  
2846 CATGCACTTTTATCTGCAATTTTGTGCTACAGAGCTCTAGAATGAAGATGAGAGTGT 2905  
2842 CATCTCTTTTGTGCTCCATTTTGCACACTGTGATATGAGAGCTCAGATGAGTGTGT 2901  
2906 TGAATTTGATGACTGTGAGAGAAATGACAGTCTGGGTCAATTTGAAGTGTGCTCA 2965  
2902 TGGCATTTGAAGATTTGTGAGAAATGAGCAAGTTTGGGTTATTTGGTGGCACTCTGCCCA 2961  
2966 CCTCTTGTGTGTTCAGAGGACTTTCAGAGTCAATAGCTGTGTATTAACCTTCAAC 3025  
2962 TCTCTTGGGAGTGTTCAGAGGCTGTGCTGAAGTGTGTGCTGGATGATCAACTTCAAC 3021  
3026 CGTGACATCAAG-----GTGAGATGATGAGAGATTTCAAGAGCTATATCAATTTCAATG 3082  
3022 AGTTAATCTCAAGGATCTGATGAGAGATGGGACTTTGTGATGATATGTTCAAGT 3081  
3083 GACTAATTTATGATCTCTTCAACCTTGTCTTATGAACTTCAATGATGATGCTGCTGC 3142  
3082 GACCAATTTGTCTACCTCCGACACAGTGTCTTGTCTATTAACCTGTGCGAATTTGGTGC 3141  
3143 TGGGTTTCAATGAGATCAATTAAGATGAAGATGAGGAGCCCTCTTTGGGAAGCT 3202  
3142 AGGAATTTCTATGCTCAATTAACAGGAGCTAACAAATCCGAGGCTCGCTTTGGAAGCT 3201  
3203 ATTTCTTGCATTTTGGGATTTGTCATCTTTATCCCTTTCAAGAGTTTGGTTGAAG 3262  
3202 GTTCTTCGATCTGGGATCTTCATCTTACCCCTTCTCAAGGATCTCATGGGAG 3261

Qy 3263 GCAAAACAGGACACCAAGATTGTCATGCTGGTCCATTCTGCTGCTTCAATCTTCTC 3322  
Db 3262 GCAGAACCGCAGCGCAACAATCGTCAATGTTGGTCCATCTCTTGGCTTATCTTCTC 3321  
Qy 3323 GCTCCTTTGGGTTGGAATTGATCCTTCTTCT 3352  
Db 3322 CTTCCTGTGGGTGAAGATCGATCCTTTCAT 3351

Search completed: December 16, 2003, 04:38:54  
Job time : 1289 secs





Db 459 GTAGAGAGAAATCAAGCTTGTCTCACTGCAAAACTCGATACAAAGATTAAAGTA 518  
Qy 424 GCCAACGTGTGACCGGTGACGAGAGAGAGACGGCGTCGATGACCTGCAACAGATTCA 483  
Db 519 GTCCACGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578  
Qy 484 ACTGAGAGCGGCTGACTGCGAGTCTGTGGCCGAGTCTGATGCTGACGGCAGATAGCT 543  
Db 579 TTGATCATGGATGAGACCTGAAACATGCCGTGAACCGGACCTCTTCAACGGCTTAA 638  
Qy 544 AGCGCGGTGAGGTGACCCCTAATGGCGCGCCCAAGAGTTTCCAGCTCAACCCCAATGTT 603  
Db 639 CCGGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683  
Qy 604 CACTCTCAACCAACGGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
Db 684 CTCTTTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743  
Qy 664 CTCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723  
Db 744 TGCCTCTTCAACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803  
Qy 724 TACTGTGCAACCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 783  
Db 804 CACTTCAACGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 863  
Qy 784 GTGTGTGTTGAGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 843  
Db 864 GTGTGTGTTGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 923  
Qy 844 CCGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876  
Db 924 TCAATTAAGCATGAG 983  
Qy 877 ATCTGATCAACCACTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936  
Db 984 ATCTGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043  
Qy 937 CATCAAGCCAGATTAATCCATATGATGATGATGATGATGATGATGATGATGATGATG 996  
Db 1044 GTTCAAGCAAAATTAATCCATATGATGATGATGATGATGATGATGATGATGATGAT 1103  
Qy 997 TCTTCTTCCACTACCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1056  
Db 1104 TTTTCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1163  
Qy 1057 CTGTATATCTGTAATCTGTTGGTTCGATGATGATGATGATGATGATGATGATGATGAT 1116  
Db 1164 CAGTTATATCGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1223  
Qy 1117 TCCCTATGAGAGAGAGACTTACCTAGACCGGCTGTCATGAGGTTTCGACAAAGAGCC 1176  
Db 1224 ATCTATATGAGAGAGAGACTTACCTAGACTCTCTCTCAGAGTACGAAAGAGAGAA 1283  
Qy 1177 AGCCATCTCAACTGCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1236  
Db 1284 AACCGTACGAGATTAGACCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1343  
Qy 1237 CTCTTTTGTGCAAAAGAAATACGTCTATCTATCTCTTGGGTGATGATGATGATGAT 1296  
Db 1344 CCCCCTTGTATACGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1403  
Qy 1297 AGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356  
Db 1404 AGGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
Qy 1357 AAACATCTGAATTTGCAAAAGAAATGAGTTCCTTTCTGCAAAACGATCAATTTGAAC 1416  
Db 1464 ATACAGCTGATTTTGTACAAAAATGAGTTCCTTTTGTAAAGATTAAATATGAGCAC 1523  
Qy 1417 GCGCTCAGAGTGTACTTCCAAACAGAAATAGCTCTTGAAGAAAGAAAGTGGCAGAA 1476  
Db 1524 GAGCTCTGAGTGTATTTTCTCGAAGAGATGATGATGATGATGATGATGATGATGATGAT 1583

Qy 1477 ACTTTGTTAGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1536  
Db 1584 CTTTGTCAAGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1643  
Qy 1537 ATGCTTATGTTCCAAAGCCGAAAGTTCCTGAAAGAGATGATGATGATGATGATGAT 1596  
Db 1644 ATGCACTGTTCTCTGCAAGAAAGTTCCTGAGAAAGTTCCTGAGAAAGTTCCTGAG 1703  
Qy 1597 CCCCCTGCTGAAACAAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1656  
Db 1704 CTCTTGGCTGAAACAAACGTCGATGATGATGATGATGATGATGATGATGATGATGAT 1763  
Qy 1657 AAAGCGAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1716  
Db 1764 ATAGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1823  
Qy 1717 AGAAAGACGAGCTATACCTATGAAAGAGTGTGATGATGATGATGATGATGATGAT 1776  
Db 1824 AGAAGCGCTGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883  
Qy 1777 TCTGTGCTGATCAAAATGCTCCATATTTGTTAACTTGGATGATGATGATGATGAT 1836  
Db 1884 TCTGTGCTGATCAAAATGCTCCATATTTGTTAACTTGGATGATGATGATGATGATGAT 1943  
Qy 1837 ACACAGCAAGCTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1896  
Db 1944 ACACAGCAAGCAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2003  
Qy 1897 AGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1956  
Db 2004 AGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063  
Qy 1957 CTAAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2016  
Db 2064 CAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2123  
Qy 2017 CAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2076  
Db 2124 CCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2183  
Qy 2077 CCAAAACAAAGAGACCATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2136  
Db 2184 CAAAGAGAGAGAGACCATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2243  
Qy 2137 GTTGTGCTTGTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2196  
Db 2244 GTTGTGCTTGTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2303  
Qy 2197 AGTTATATTTTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2256  
Db 2304 AGACTTCA-----AGCAGATTCATGCGCTAGAGAAATGTCAGAGAG 2345  
Qy 2257 AGGAGCTGCTCAAGAGCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2316  
Db 2346 GTGTATGCTCCAGAGTGAATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2405  
Qy 2317 AAAAGAAATTTGAGCAATCTTCTGTTTGTGATACACACTTCTCGAAGAGTGGAA 2376  
Db 2406 AGAAGAGTTTGAACATCTCCGTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2465  
Qy 2377 CTTGAGAGTCAAGTCTGCTTCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2436  
Db 2466 TTCCCGCTAAGCAGAGCCCGAGTGTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2525  
Qy 2437 GTTATGAAG 2496  
Db 2526 GGTACCAAGATTAACGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2585  
Qy 2497 AAGATATTTAACTGTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2556  
Db 2586 AAGATATCTGACGGTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2645

QY	2557	CTAAACGGGTTCATTTCAAAAGGTCGTGACCCCTGTAAATCTTTGACATGCGTCTTCCACGAG	261
Db	2646	CTTAACCGTGACGCTTTTAAAGGATCTGCTCTATTAACTGTGCAGATGGCTTCAATCAAG	2705
QY	2617	TGCTTCGGGTGGGCTCTGGGTCCTATTGAGATCTTCTTGAGCAATCATTTGCCCTTTTGGT	2676
Db	2706	TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTTGTCCGATNTGGT	2765
QY	2677	ATGGGTATGTTGGCGGCTCTGAAATTTTGGAAAGATTTTCTTACATCAACTCCATCGTGT	2736
Db	2766	ATGGTTATGTTGGTGGTTTAAATAGTTGGAGAGATCTCTTACATCAACTCTGTGCTCT	2825
QY	2737	ATCCCTGGACATCTAATTTCCCTCTTGGCTTACTGTACATATGCTGGCAATCTGTTTATTTGA	2798
Db	2826	ATCCCTGGACCTTCACTTCCATTGATCGTCTATTGTCTCTCCCGGGTTTGTATTCACA	2885
QY	2797	CAGGGAATTTTATCACTCCAGAGCTGATTAATGTTGGCCAGCGCTGGTTCATGTCACCTT	2856
Db	2886	CAGGAAATTTATCTCTCTCCCTGAGATAGACACTACGCAAGTATCTTCAATGCTCATGT	2945
QY	2857	TTATCTGCAATTTTTCCTACAGACATCTTGAATAGATGAGATGGATGGTGGAAATTTGAT	2916
Db	2946	TCAATATTCATAGCAGTAACTGGAAATCTTCGAAATGCAATGGGAGAGTGTCCGAATCAGT	3005
QY	2917	ACTGTTGAGGAATAGACAGATTTCTGGGTCATTGAGAGTGTGTCTCTCAACCTCTTGTCTG	2976
Db	3006	ATGTGTGAGAAACAGACAGTTTGGGTAAATCGGAGGGGCGCTCTCCGATCTAATTTGGCTC	3065
QY	2977	TGTTCCAGGACCTTCTCAAGGTCATAGCGGTGTGATTAACAAGCTTCAACCGTCAATCAA	3036
Db	3066	TGTTTCAAGGTTTGTCTCAAGTTTCTAGCCGAGTTTACACGAAATTTCAAGTCACTTCAA	3125
QY	3037	AGGGTGAAGATGATAGAGAGTTCTTCAGAGCTAATATACATTCAATGACACTACTTATTTGA	3096
Db	3126	AAGCAGCAGACGATGAGACTTCTCTGAGCTTTAATCTTCAAGTGACAACTTTGTTGA	3185
QY	3097	TACCTCTTACCACTTGTCTTATTGAACTTCATTTGGTGTGTGCTGGCGCTTTCMAATG	3156
Db	3186	TTCCCTCCACAACACTTCTGATCATTAATCATATGAGTTAATGTCCGGCTTTCTGATG	3245
QY	3157	CGATCAATTAAGGATGATAGTACATGAGGAGCCCTCTTTGGGAAGCTAATCTTTGCATTTT	3216
Db	3246	CCATTAGCAATGGCATATACATCATGAGGAGCCCTCTTTGGGAGACTTTTCTGCTCTTT	3305
QY	3217	GGGTGATTTGCCATCTTATCCCTTTCTCAAGGTTTGGTTGGAAAGCAAAACAGACAC	3276
Db	3306	GGGTCAATTTGTTATTAATCAATCTCTCAAGGAAATCTTGGGAGCAACAAATATGC	3355
QY	3277	CAAGCATTTGCATCGTCTGTTCATTTGCTGTGGCTTCAATCTTCCGCTCTTTGGGTTCC	3336
Db	3366	CTACGATTAATGTGGTGTGTCTAATTTCTTAGCTTCATCTTGACATCTTTGTGGGTTCA	3425
QY	3337	GGATTGATCCCTTTCCTTCCGAAGATATGTCGCTTCTTGAGAGAGTGTGGTTGGATT	3396
Db	3426	GAATTTAACCCGTTGTGTGCTA---AAGGGGACCAAGTGTGGAGATCTGTGTCTGAATT	3482
QY	3397		
Db	3483		

```

; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU PO0699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3673
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)..(3313)
US-09-221-013A-11

Query Match      33.8%; Score 1330.8; DB 4; Length 3673;
Best Local Similarity 64.9%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 1067; Indels 75; Gaps 7

QY    127 GTGAGGGGAGGAGACGATGTGAGGGAGCAGCGCCGGGCTGTGGGCCGCCTCCCAACAACCGA   186
DB     54 GTGTTCGGTGTGCTCGCATGTGAGGSCCATGTCGGCTTGTTGGTCTGAATTCTTAACGGAGA   113

QY    187 ACGACTCTGTGCTCATTCGCCCGCACGGCGATCCCGGGCCGAAAGCCCAGCGAGCAGA   246
DB     114 ACGAACCTGTTCGATCCGATCGCATGTGAATTGTATGGCGGAGCCAAACCTTTGAAGAAATATGA   173

QY    247 ACGGCAGGTGTGCCAATTTTGGCCGACGACGTGCGCTTGCCCCCGGCGGAGCCCTT   306
DB     174 ATGGCCAATATGTAGATCTGTGTATGATGTTGATCTGCTGMAACTGGAGATGTCT   233

QY    307 TCGTAGGCTGTCAAACGAGTGGCCCTTCCCGTCTGGCCGGGACGTGTAAGATAACGAGCC   366
DB     234 TTGTGCGGTGTAATAATGTGCTTCCCTGTGTGTGCGGCTTGTCTATAGTACGAGAGA   293

QY    367 GGGAGGAGCAACGCAAACTGCCCCAGTGCAAGACTCGATACAGGCGCTCAAGGCGTCC   426
DB     294 AAGATGGAATCTCAGTGTGGCCCTCAATGCAAGACTAATTTAGACGACACAGGGGGAGTC   353

QY    427 AACGTGTACCCGGTGAACGAGAGAGAGCGCGCTGATGACTTGACACAAGTTCAACT   486
DB     354 CTCGTGTGAAGAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATCA---   410

QY    487 GGGAAGGCGCATGACTCGGAGTGTGGCGCGAGTGCATCTGTACGACCAATGAGTACG   546
DB     411 -----ATTAGCCCCAGGAGGTAAACAAGCGAGACACACACGCCA---TGGCGAAG   458

QY    547 GCCGTGAGGTGACCCCTAATAGGCGCGCCACAAGCTTTCAAGTCAACCCCAATGTTCCAC   606
DB     459 AGTTTTCTTTCTTCCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT   518

QY    607 TCCTCACAACGCGGCAAAATGATGATGATCAATCCACCGAGACAGACGCGTGTGCTT   666
DB     519 CGGTTTCTGAGAGATTTGCAAGCGCTGTATCACAATCTGTGCAACTGATCAAGGCTCTT   578

QY    667 CTTCATGAGGTGGTGGGGAAAAGAGATACATCCCTCTTCTATTCGGATCCCAAGTTCAC   726
DB     579 -----TGGGTCTCTTGACAGAAATGCTATTTATCTTCATATATATGATCCACGGCAAC   632

QY    727 CTGTGCAACCCAGGCTATATGACCCATCCAGAGATCTTGTCGATATGGATATGATGATG   786
DB     633 CTGTCCCTGTAAGAAATCGTGAACCCGCTCAAAAGACTGAATCTTTATGGGCTGTGTAATG   692

QY    787 TTGCTTGGAGAAACGATGAGAAATTGGAAGCAGAGACAAAGAGAGATG-----   836
DB     693 TTGACTGAAAAAGAAAGAGTTGAAGGCTGGAAGCTGAAGCAGAGAAAAATATGTTACAGA   752

QY    837 -----CACCAAGCGGGAAATGATGCGTGTGTGTATGATGATGATGATGATGATG   879
DB     753 TGACTGTAAATACCATGAAAGGAAAAGGAGAGAAATTTGAAGGAGCTGTGTTCCATGCG   812

QY    880 CTGATCTACCACTATGATGATGACCAAGACAAACAACATGTCTCAGGAAAAATTTCCATTCAT   939

```

Db 813 AAGAACTCCAAATGGCTGATGATACAGCTTCTCTATGAGTGGTGGCTATCCAT 872  
Qy 940 CAAGCAGATTAATCCATATAGATATATCATATATGCGGTGGTGGTGGTGGTCT 999  
Db 873 CTTCTGCGCTAACCCCTTATCCGGTGTGATTAATTCGCGCTTATCATCTGTGTCT 932  
Qy 1000 TCTTCCATACCGAGTATGATCCGGTGAATGATCATTTGGTTGGCTCATATCTG 1059  
Db 933 TCTTGCATATATGATCAATCACTCCCTGAAAAATCATATCTTGTGTGAGCTCCG 992  
Qy 1060 TATCTGTAATTCGCTTGGCAATGCTGATTCCTTGAATCAATCCCAAGTGTCTC 1119  
Db 993 TATCTGTAATTCGCTTGGCAATGCTGATTCCTTGAATCAATCCCAAGTGTCTC 1052  
Qy 1120 CTATTGAGAGAGAGATTAATCTAGACGGCTGTCACTGAGTTGCAAGAAGGCGAC 1179  
Db 1053 CCATTAACAGGAGAGATTAATCTAGACGGCTGTCACTGAGTTGCAAGAAGGCGAC 1112  
Qy 1180 CATCTCACTGCTCCAAATGATTTCTTTGTCAATGAGTGTGATCCCTTAAAGAACTC 1239  
Db 1113 CATCAAGCTCGTTCCTGTTGATGTTGTTGATGATCAGTGGACCATTTGAAAGGCTC 1172  
Qy 1240 CTTTGGTCACAAAATATCTGTTCTATCTATCCTTCCGGTGAATTAATCCTGTTGAAG 1299  
Db 1173 CCGTTGTACAGAAACAGATTCCTCAATCTTCTGTGATCAACCGATGATTAAG 1232  
Qy 1300 TTTCTGTATGTTTGTGATGATGCTGCAATGCTAACGTTTGAAGATTAATCGAA 1359  
Db 1233 TAGCCGTATATGTTTGAATGATGATGCTGCAATGCTAACGTTTGAAGATTAATCG 1292  
Qy 1360 CATCTGAATTTGCAAGAAATGGTTCCTTTCTGCAACGGTACAAATTTGAACCTCGC 1419  
Db 1293 CCGCTAGTTTGGCAAGAAATGGTTCCTTTCTGCAAGAAATTTGAACCTGAGG 1352  
Qy 1420 CTCAGAGTGTACTTCCAAAGATATGATCTTGAAGAAAGAGTGGCGAGCAACT 1479  
Db 1353 CCGCTGAATTTCTATTTTGGCCAGAAATGATTAATCTTGAAGAAAGAGTGGCGAG 1412  
Qy 1480 TTGTTAGGAGAGAGAGATTAATGAGAGATGATGAGAAATTCAGAGTGAATCAATG 1539  
Db 1413 TTGTTAAAGAGAGAGAGATTAATGAGAGATGATGAGAAATTCAGAGTGAATCAATG 1472  
Qy 1540 CCTTATGTCCAAAAGCCAGAAAGTTCCTGAAGAAAGATGAGCAATGCAAGATGAAACC 1599  
Db 1473 CTCTTGTGCAAAAGCAAGAAATTCCTGAAGAAAGTTCCTGAAGAAAGATGAAACC 1532  
Qy 1600 CCGTGGCTGGAACCAATGTTCTGATCATCTCGGAATGATTAATGAGTCTTCTTGGCCAA 1659  
Db 1533 CCGTGGCTGGAACCAATGTTCTGATCATCTCGGAATGATTAATGAGTCTTCTTGGCC 1592  
Qy 1660 GCGAGGCTGTAAGTGTGAGGAGAAATGAATCGCAAGTGTGTTATGTTTCTAGAGAG 1719  
Db 1593 GTGGGGGTGTGATACCGATGAGAAATGAGTGTGCTGATCATCTATGTTTCTCGGAAA 1652  
Qy 1720 AACGACAGGCTTAATCAATCATATGAGAAAGCTGTGCTATGAATGCAATGCTCGAGTCT 1779  
Db 1653 AGGGGCTGGAATTTCAACACCAAAAGGCTGGAATGATGATGATGATGATGATGAT 1712  
Qy 1780 CTGCTGTACTAACAATGCTCCATATTTGTTAACTTGAATGATGATGATGATGATGAT 1839  
Db 1713 CTGTTGTTCTTACCAATGAGATATCTTTGAACCTGATGATGATGATGATGATGAT 1772  
Qy 1840 ACAGCAAGGCTTAATGAGAAAGATGTTTATGATGAGCCCTTACTAGAGAAAGAG 1899  
Db 1773 ACAGTAAGGCTTAATGAGAAAGATGTTTATGATGAGCCCGCTTACTAGAGAAAGAT 1832  
Qy 1900 TTGCTATGATACAGTTCCCTCAAAAGATTTGATGAGTTGATGCGCATGACCATATGCTA 1959  
Db 1833 GGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1892  
Qy 1960 ACCGGAATGTTGCTTTTGTATATCAATGAAAGTTTGAATGATGATGATGATGATGAT 2019

Db 1893 ACAGGAATATATGCTTTTTCATATTAATCAATGAAAGGCTGTGATGATGATGATGATGAT 1952  
Qy 2020 TTTATGTTGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2079  
Db 1953 TATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2012  
Qy 2080 AAACAAAGAGGCAATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2139  
Db 2013 TGACGGAAGAAATTTTAAACCAATATTA-----TTGTCAAGA 2051  
Qy 2140 GCTGCTTGGCAATTTGAGAGCAAAAGAGATCAACCAACCAAGAGAGAGAGAGAGAGAGAG 2199  
Db 2032 GCTGCTGCGGCTCAAG 2108  
Qy 2200 TATATTTTTCAG 2259  
Db 2109 GAGAGAGCATCAAG 2168  
Qy 2260 AAGCTGCTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2319  
Db 2169 AAGGTTTGAAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2228  
Qy 2320 AGAAATTTGGCAATCTTCTGTTTGTATCATCACTCTGAGAAATGATGAGAGAGAGAG 2379  
Db 2229 AGCTTTTGGTCAAGTGGCTGATTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2288  
Qy 2380 TGAAGAGTCAAGTCTGCTCTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2439  
Db 2289 CACCAACCAACCAATCCGCTACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2348  
Qy 2440 ATGAAGACAG 2499  
Db 2349 ACAGAGACAG 2408  
Qy 2500 ATATCTTAATGTTTCAAG 2559  
Db 2409 ATATCTTAATGTTTCAAG 2468  
Qy 2560 AACGGTTGCAATTCAG 2619  
Db 2469 CACGCTCGGCTTCAAG 2528  
Qy 2620 TTGGTGGGCTCTTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2679  
Db 2529 TTGATGAGGCTTTGGGATCTATGAGATCTTCTTGAAGAGAGAGAGAGAGAGAGAGAG 2588  
Qy 2680 GGTATGAGGAGGAGTGAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2739  
Db 2589 GTTACATGAG 2648  
Qy 2740 CTTGACATCTATTTCCCTCTTGGCTTATGATGATGATGATGATGATGATGATGATGATGAT 2799  
Db 2649 CTATTAATCAATCCCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2708  
Qy 2800 GGAATTTTATCACTCAAG 2859  
Db 2709 ACAGATTCATATACCAAG 2768  
Qy 2860 TCTGATTTTGTCTAG 2919  
Db 2769 TCTCAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2828  
Qy 2920 GGTGAGAGATGAG 2979  
Db 2829 GGTGAG 2888  
Qy 2980 TCCAGGAGCTTCTCAAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3039  
Db 2889 TCCAGGAGCTTCTCAAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2948  
Qy 3040 ---GTGAG 3096  
Db 2949 CCAAG 3008

Accession	Sequence	Position
QY	3097 TACCTCCACCAACCTTGCTGTTCTATTGAACCTTCATTGGTGTGTGCTGGCTTCAATG	3156
Db	3009 TTCAACCAACACCGCTCTACTCTTGAACCTTCATAGGACATTGTGTGTGTCTCTCTTATG	3068
QY	3157 GCATCAATTAACGGATATAGTCATGAGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTT	3216
Db	3069 CTGTAAACAGTGGCTACCAAGTCGTGTGGGTCGGCTTTTGGGAAGCTCTTCTTGCGTTAT	3128
QY	3217 GGGGATTTGTCCATCTTTATCCCTTTCTCAAAAGTTTGTGTGAAGGCAAAACAGACAC	3276
Db	3129 GGGTATTGGCCATCTCAACCTTTCTTGAAGGCTGTGTGGAAAGCAAAACCGAATAC	3188
QY	3277 CAACGATTGCATCGCTGTGCATCTGTGTGAGCTTCAATCTCTGCTCTTGGGTTTC	3336
Db	3189 CAACCATGTGATTTGTGTGTGTCTTCTTGCGCTCATCTTCTCGTTGTGGGTCA	3248
QY	3337 GGATTGATCTTTCTCTTG	3354
Db	3249 GGATCAATCCCTTTGTGG	3266

### RESULT 3

US-09-221-013A-5

; Sequence 5, Application US/09221013A

; Patent No. 6495740

; GENERAL INFORMATION:

APPLICANT: Arioli, Antonio

APPLICANT: Williamson, Richard E.

APPLICANT: Betzner, Andreas S.

APPLICANT: Peng,

TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan

FILE REFERENCE:

CURRENT APPLICATION NUMBER: 05/09/221, 013A

PRIOR ADDITION ATTEMPTED: DCE/2/1007/004003  
CURRENT FILING DATE: 1998-12-23

PRIOR FILING DATE: 1997-05-24

PRIOR APPLICATION NUMBER: AU 800698

PRIOR FILING DATE: 1996-06-27

NUMBER OF SEQ ID NOS: 37

SOFTWARE: Patent In Ver. 2.0

! SEO ID NO 5

LENGTH: 3603

TYPE: DNA

ORGANISM: *Arabidopsis thaliana*

**FEATURE:**

NAME/KEY: CD:

; LOCATION: (1)..(3243)

US-09-221-013A-5

Query Match	33.7%	Score	1326.6	DB	4	Length	3603
Best Local Similarity	65.0%	Pred.	No.	0			
Matches 2107; Conservative		Mismatches	1059;	IndeIs		Gaps	71

OY	144	ATGAGGCGAGGCGCCGGGCTGTGTGGCCGGCTCCCAACCGCAACGAGTCTGTCTATC	203
Db	1	ATGAGGGCCAGTGGCCGGCTTGTGTGTCTGATCTCAACGGAGAAAGAGCTCTGTGGATC	60
OY	204	CGCCCGCAGCGCGGATCTCCGGGCCGAAGCGCGCGCGAGACAAACGGGCGAGGTGTGCAG	263
Db	61	CGACATGATCTGATGTGGCGGACCAAACTTTGAAGAAATGTAAATGGCCAGATATGTCAAG	120
OY	264	ATTTCGGCGACGACGTCGTGGCTTGTGCCCCCGGGCGGACCCCTTGTGTGCGTGCACGAG	323
Db	121	ATCTGTGTGTGATGATGTGTGATCTGCCTGAAACTGGAGATGTCTTTGTGCCTGTATATGAA	180
OY	324	TGCGGCTTCCCGCTGTGCGGGGACGTCTACGAAATACGAGCGCGGGAGAGGGCACGACGAC	383
Db	181	TGTGCTTCCCTGTGTGTGTGCGCTTGTCTATGTAGTACGAGGAAAGATGGAACTCAGTGT	240
OY	384	TGCCCCCAGTGCAGAGCTCGATACAGCGGCTTCAAGGGCTGCGCAACGTGTGACCGGTAC	443
Db	241	TGCCCTCATGTACAACTAGATTTCAGACGACACAGGGGGAGTCTCTGTGTGTAAAGAAAT	300

QY	444	GAGGAGGAGGACGGCGCTCGATGACTGAGCAACAGAGTTCAACTGCGGACGGCATGACTCG	503
Db	301	GAAGATGAGGATGATGTTGATGATATGAGAAATGAGTTCA-----ATTAGGCC	348
QY	504	CAGTCTGTGGCGGAGTCCATGCTCTTACGGGCACTAGAGCTACGGGCGGTGAGGTGACCT	563
Db	349	CAGGGAGCTACAAAGGGGAGACACCAACGCCA---TGGCGAAGAGTTTCTTCTTCTCT	405
QY	564	AATGGCGGCCCAAGCTTTCAGCTCAACCCCCAATGTTTCACTCTCACCAACGGGCAA	623
Db	406	AGACATGAATCTCAACCAATTCCTCTTCTCAACCATGGCAATACGAGTTTCTGAGAGATT	465
QY	624	ATGATGATGATGATCAATCCACGGAGACACGCGGTGTGCTTCTTTCATGGGTGGGG	683
Db	466	CGACGCGCTGATACAAATCTGTGCGAATCAATCAAGTCTCTT-----TGGGTCCTTCT	519
QY	684	GGAAGAGGATATACATCCCTCTCTTATGCGGATCCACAGTTTACCTGTGCAACCCAGGCT	743
Db	520	GACAGGAATGCTATTTATCTCCATATATGATTCACAGGCAACCTGTCCCTGTAAAGATC	579
QY	744	ATGAGCCCATTCAAAGATCTTGTCTGCATATGGGTATGTATGTGTCTTGGAGAAACGG	803
Db	580	GTGACCCCGTCAAAAGACTTGAACCTTATAGGCGCTTGTAATGTGTGATGAGAAAGAAAG	639
QY	804	ATGAGAAATTGGAAAGCAGAGCAAGAGAGATGC-----	837
Db	640	GTTAAAGCTGTGAAGCTGTGAAGGAGAGAAAATATGTACAGATGACTGTGTAATACAT	699
QY	838	-ACCAAGCGGGAAATGATGTTGTGTGATGATAGTGAACAGTGTGATCTTACCACTATG	896
Db	700	GAAAGGAAAGGAGAGAAATTGAAGGAGCTGTTCCAAATGGCGAAGAACTCCAAATGGCT	759
QY	897	GATTAAGCAAGACAAACACTGTCCAGGAAAAATTCATTCTTCATCAAGCCAGATTAATCCA	956
Db	760	GATATACAGCTCTTCTTAATGAGTGTGTGGTGTCTTATCCATCTTCTGCGCTTAACCCCT	819
QY	957	TATAGATGATATCATTTATTCGGCTGTGGTTTGGGGTCTTCTTCCACTACCGAGTG	1016
Db	820	TATCGGGTGTGATTAATTCGCCGCTTATCATCTGTGTCTTCTTGGCAATATCGTACA	879
QY	1017	ATGATATCCGAGTAATGATGCTTTGCTTGTGTGCTCATATCTGTTATCTGTAAATCTGG	1076
Db	880	ACTCAACCTGTGAAGAAATGCAATATCCCTTGTGTGTGACCTCGGTTATCTGTGAGATCTGG	939
QY	1077	TTTGCATATGCTTGTGATCATTTCCAAATGCCAAAGGTTCCCTATTGAGAGAACT	1136
Db	940	TTTGCATTTTCTTGGCTTCTTGTACATGTTTCCAAATGTATCCCATTTAACAGGGAATCT	999
QY	1137	TACCTAGAACGGCTGTCACTGAGGTTTGCAGACAGAAAGCCAGACCATCTCAACTTGTCCA	1196
Db	1000	TATCTTACGCGTCCGCTATTAAGTATGATCGAGACGCTGAACATCAAGCTGTGCTCT	1059
QY	1197	ATTATATTTCTTGTCAATGAGTTGATATCCCTTAAAGAACTCTTTGTGTCAACAAT	1256
Db	1060	GTTATATGTGTTGTATATACAGTGAACCAATGMAAGAGCTCCCTTGTATACGCAAC	1119
QY	1257	ACTGTTCTATCTATCCTTTCGGTGGATTAATTCGTGTGATTAAGGTTCTTGTCTAATGTTCT	1316
Db	1120	ACATTTCTCTGATTTCTTTCTGTGACCTACCGGTATTAAGTATAGCTGTATATGTTTCA	1179
QY	1317	GATGATGCTGTGCATATGCTTAAAGCTTAAAGCATTAATCTGAACATCTGAATTTGCAAG	1376
Db	1180	GATATATGTTCAAGTAAATGCTTAACTTTGAAATCCCTTCTGAACCGCTGAGTTTGGCAAG	1239
QY	1377	AAATGGGTTCTTTCTGTCAAAACGGTACAAATATTGAACCTGCGCGTCCAGAGTGTATCTTC	1436
Db	1240	AAATGGGTACATTTTGTGCAAGAAATTAACATTTGAACCTTGAAGGCCCTGAAATTTCTATTTT	1299
QY	1437	CAACAGAAATGATGACTTGAAGACAAAGTGGCAGCAACTTGTATGGAGAGAGGA	1496
Db	1300	GCCGCAAGAAATATGATTTCTTGAAGAGCAAAATCCAAACGCTTTTGTATTAAGGACGCA	1359

QY 1497 GCAATGAGAGAGATGAGAAATTCAGAGTGAAGATCAATGCCCTTAGTGGCAAGCC 1556  
Db 1360 GCTATGAGAGAGATGAGAAAGTTAAAGTGAAGATAAATGCTCTTGTGGCAAGCA 1419  
QY 1557 CAGAAAGTTCCTGAAGAGATGAGACATGACAGATGAGAAACCCCTGGCTGGAAACAAT 1616  
Db 1420 CAGAAATCCCTGAAGAGATGAGACATGAGAGATGAGATCCCTGGCTGGAAACAAC 1479  
QY 1617 GTTCGTGATCATCTGGAATGATTCAGAGTCTTCCTGGCCMAAGCCGAGGCCCTGA 1676  
Db 1480 ACTAGAGATCATCTGGAATGATGACAGTGTCTTAGGCCATAGTGGGGGTCTGATACC 1539  
QY 1677 GAGGAAATGAAATCTGCCAGATGTTGTTATGTTCTAGAGAGAAAGACAGGCTTAAC 1736  
Db 1540 GATGAGAAATGAGTGGCTGACATCACTATGTTCTGTGMAAAGGCGCTGATTTCAA 1599  
QY 1737 CATCATAGAAAGCTGCTGATGAAATGATGATGATGATGATGATGATGATGATGAT 1796  
Db 1600 CACCAAAAAGAGCTGAGCTATGAAATGATGATGATGATGATGATGATGATGAT 1659  
QY 1797 GCTCATATTTGTTAAACCTTGATGATGATGATGATGATGATGATGATGATGAT 1856  
Db 1660 GAGCATATCTTTGAAACGATGATGATGATGATGATGATGATGATGATGATGAT 1719  
QY 1857 GAAAGCATGCTTTTATGATGAGACCCCTTACTAGAGAAAGAGTTGCTATGACAGTTC 1916  
Db 1720 GAGGCTATGCTTTATGATGAGACCCGCTATGAGAAAGAGTGTGCTATGCTCAGTTC 1779  
QY 1917 CCTCAAGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1976  
Db 1780 CCTCAAGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839  
QY 1977 TTGATATCAACATGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 2036  
Db 1840 TTGATATCAACATGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1899  
QY 2037 TGTGATTTGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2096  
Db 1900 TGTGATTTGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1959  
QY 2097 TCAAGACTGCAACTGCTGAGCCAGTGTGCTTTGCTGCTGCTGCTGCTGCTGCTGCT 2156  
Db 1960 GAACCAATATTA-----TTGTCAAGAGCTGTGTGCGGCTCAAGG 1998  
QY 2157 AAGCAAAAGAGACTACCAAAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2216  
Db 1999 AAGAAAGTAAAGTAGCA---GAAGTATACAGAGAGAGAGAGAGAGAGAGAGAGAG 2055  
QY 2217 GAAAGAGACCAATCCCTGATATGCTCTGTGGAATGACAGAGCTGCTCAGAGACT 2276  
Db 2056 AGTGACTCCAAATGCTCCACTTTCAATATGAGAGACATCGATAGAGGCTTTTGAAGGTTAT 2115  
QY 2277 GAGATGAAAGGCGGCTATGTAATCAACAAAAATTAGAAAGAAATTTGGCAATCT 2336  
Db 2116 GATGATGAGAGCTTATTTCTAATGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2175  
QY 2337 TCTGTTTTTGTATCATCCACTCTGAGATGATGAGAACTTTGAAGAGTGCAGTCT 2396  
Db 2176 CCGGTATTTATGCGGCAACCTTCATGAGCAAGGCGGCAATCCACCAACCAATCC 2235  
QY 2397 GCTTCTTTTGAAGAGAGCTATACATGATGATGATGATGATGATGATGATGATGAT 2456  
Db 2236 GCTACTCTTTGAAGAGAGCTATGATGATGATGATGATGATGATGATGATGATGAT 2295  
QY 2457 TGGGAAAAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2516  
Db 2296 TGGGCAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355  
QY 2517 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2576  
Db 2356 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415  
QY 2577 GGTTCGACCTCTGAATCTTTCAAGATGCTTTCACCAAGGCTGCTCGGTGGGCTTTGGG 2636

Db 2416 GATCTGACCAATCATCTTTCTGATGCTTTGAACCAAGTTCTTGCATGAGGCTTTGGGA 2475  
QY 2637 TCTATGAGATCTTCTGAGCAATCATGATGATGATGATGATGATGATGATGATGATGAT 2696  
Db 2476 TCTATGAGATCTTCTGAGCAATCATGATGATGATGATGATGATGATGATGATGATGAT 2535  
QY 2697 AATTTTTGAAAGATTTTCTTCAATCAATCTCATGATGATGATGATGATGATGATGAT 2756  
Db 2536 AGACTTTTGAAGAGATGCTTATGATCAACACATGCTCTATTCATTAATCAATCCCT 2595  
QY 2757 CTCTGGCTTCTGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2816  
Db 2596 CTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2655  
QY 2817 GAGCTGAATTAATGTTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2876  
Db 2656 GAGATAGAGACTAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2715  
QY 2877 AGCATCTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2936  
Db 2716 GGAATCTGAGAGCTGAGATGAGAGCGGTGAGATGAGATGAGATGAGATGAGATGAG 2775  
QY 2937 TTCTGGGCTCATGAGAGGCTGCTCAGACCTCTTGTGCTGCTGCTGCTGCTGCTGCT 2996  
Db 2776 TTCTGGGCTCATGAGAGGCTGCTCAGACCTCTTGTGCTGCTGCTGCTGCTGCTGCT 2835  
QY 2997 GTCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3053  
Db 2836 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2895  
QY 3054 GAGTCTCAGAGCTATATACATTAATGAGACTACCTTATGATGATGATGATGATGAT 3113  
Db 2896 GATTTTGAAGACTATACATCTTCAATGAGAGCTCTTCTATCTTCAACCAACAGCTC 2955  
QY 3114 CTTTATGAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3173  
Db 2956 CTACTTGAAGCTTATGAG 3015  
QY 3174 GAGTCATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3233  
Db 3016 CAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3075  
QY 3234 TATCCCTTTTCAAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293  
Db 3076 TACCCTTTTCAAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3135  
QY 3294 TGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3353  
Db 3136 TGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3195  
QY 3354 G 3354  
Db 3196 G 3196

RESULT 4  
US-09-221-013A-9  
; Sequence 9, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27

NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 3614  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURES:  
NAME/KEY: CDS  
LOCATION: (217)..(3411)  
US-09-221-013A-9

Query Match 32.2%; Score 1269.2; DB 4; Length 3614;  
Best Local Similarity 63.8%; Pred No. 0;  
Matches 2051; Conservative 0; Mismatches 1108; Indels 57; Gaps 6;

QY 209 CGACGCGATCCCGGCGGAGACCGCGGAGAGAGACGCGAGGTGTGCGAGATTG 268  
DB 225 CGAAGAGAAACCGCGGAGACCGATGAAACATTGTCCGAGACTTGGCAGATCTG 284  
QY 269 CGGCGACGACGTGGCCTTGGCCCCGCGGAGACCCCTTGTGGGTGCAAGATGCGC 328  
DB 285 TAGTGACATGTGTGCAAGACTGTGATGAGATGTTTGTGGCTGTGATTTGTTTC 344  
QY 329 CTCCCGCTGCGGAGACTGTAAGAAATAGAGGCGGAGGAGGAGCGAGACTGGCC 388  
DB 345 ATCCAGTTGTGCGCCCTTGCTACGATATGAGAGAAAGATGGAAATCAATCTTGTCC 404  
QY 389 CCAATGCAAGACTGCATACAGCGCTCAAGGCTGCCAACGTGTGACGAGAGAG 448  
DB 405 TCAGTGCAGAAACCAATACAGAGGCTCAAGGTAGTCTGTATCTTCTGTGATTAAGA 464  
QY 449 GAGAGCGCGCTGTGATGACTGAGCAAGATTCACTGGAGCGGCAATGACTGCGAGTC 508  
DB 465 CGAGGATGCTTGAAGTGAAGTACTG---TTGAGTTCAACCTCAGAGAGAGAA 521  
QY 509 TGTGGCGAGTCCATGCTCTACGCGCAATGAGCTACGCGCGTGGAGGATCCCTAATGG 568  
DB 522 AATTTCAAGCGGAGTGTGCTGCTGCTCACTTCACTGCGGAGAGGAGAAATGGGAGA 581  
QY 569 CGCGCCAGAGCTTTCAGACTCAACCCCA---ATGTCCACTCTCAACCAAGGCAAT 625  
DB 582 ACCCAGATGATTAAGAGGTCTCAACATCATCTTCTCGTCTCAAGAGAGAGAGAA 641  
QY 626 GGTGATGATCCACCGGAGAGAGCGCGTGTGCTTCTTCAATGAGTGTGGGCG 685  
DB 642 TACTTCAGAGATTTTCTGCTGCTCACTGAAGGCTCTGTATCTTCAATATGCGC 701  
QY 686 AAGAGATACATCCCTTCTTATAGCGATCCAGCTTACTGTGCAACCGAGTCTAT 745  
DB 702 TGGGGAGAAAGCGCTTCCCTATTCATCAGATGTCATCAATCAACAAATAGAAAGATTGT 761  
QY 746 GAACCATCCAGATCTTGTGCAATAGGATGATGATGTTGCTTGGAGAGAGAGAT 805  
DB 762 GATCTCTGTGAGCTCGGAGATGTAGCTTGAAGAGAGATTTGCTGAGAGAAATGA 821  
QY 806 GAGAGATTGAGAGAGAGAGAGAGAGATGACAGAGCGGAGAAATGATGTGTGTGA 865  
DB 822 GGAAGAGAAATATCTGGTCTGTGAGAGCGAGGCTGTTGAAAAGGTGGAGTGA 881  
QY 866 TGATGTGATGCTGTATCA---CTAATGATGAGAGAGAGAGAGAGAGAGAGAG 910  
DB 882 TATTATGTCAGACAGATATCTAGCAGATGAGGCTGTGAAATGAGAGAGAGAGAG 941  
QY 911 ACAATGTCAGAGAAATTCACCTTCATCAAGCAGATTAATTCATATAGATGATTTAT 970  
DB 942 GCTTCTGTCAAGAAAGTTTCAATTCCTTCAACAGATCAATCTTACAGATGATTTAT 1001  
QY 971 CATTAATGCGCTTGTGTGTGGGTCTTCTTCACTACAGAGATGATGATCGGATGA 1030  
DB 1002 TATGCTGGGGCTTGTATCTTGTCTTCTTGTGATTAACGATTAACAAACCGAGTGC 1061  
QY 1031 TGATGATTTGCTTGTGGCTCATATCTGTATCTGTGAATCTGTGTTGCAATGCTTTG 1090

DB 1062 AATGCTTGTCTATAGGCTGTCTGTGATATGATGATCTGTTGCTTATCTGT 1121  
QY 1091 GATCTGTATTAATCCCAAGGTGCTCCCTATTTAGAGAGAGACTTAATCTAGACGGCT 1150  
DB 1122 GATTTGATTCAGTTTCCCAAGGTGTTCCGTGAAACCGTGAACCTACCTCGACAGGCT 1181  
QY 1151 GTCACTGAGTTGAGAGAGAGGCGAGGCTCACTCACTGCTCCCAATTTGATTTCTTGT 1210  
DB 1182 TGCTTTAAGATATATGCTGAGAGGTGAGCCATCAAGTTAGCTGTGTCATTTTGT 1241  
QY 1211 CAGTACGTTGATCCCTTAAAGAACTCTTGTGTGACAAACAATCTGTCTATCTAT 1270  
DB 1242 GAGTACGTGTAACCCCTTGAAGAGAGCCACCCCTTGTGACAGCCAAACAGTCTCTAT 1301  
QY 1271 CTTTCCGTGATTAATCTGTGTAAGGTTCTTGTCTATGTTTGTGATGATGTTGCTGC 1330  
DB 1302 TCTGCTGTGACATACAGGAGTGAAGAGTGTCTGTATATTTCTGATGATGTTGCTGC 1361  
QY 1331 AATGTAACGTTTGAAGATTAATCTGAACATCTGAATTTGCAAGAAATGGGTTCTT 1390  
DB 1362 TATGTTATCAATTAATCACTGCAAGAAACATCAAGTTGCTGTAAATGGTATCAAT 1421  
QY 1391 CTGCAACGCTACATATTTGAACCTCGCTCCAGAGTGTACTTCCAAAGAGATAGA 1450  
DB 1422 TTGCAAGAAATATAGCATAGAGCTCGTGACAGAGATGTACTTTGCTGGAATAAGA 1481  
QY 1451 CTACTTGAAGACAGGTGAGCAAACTTTGTAGAGAGAGAGACATAGAGAGA 1510  
DB 1482 TTACTTGAAGATTAAGTTTGAAGATCATTTGTCAAGATGTAGAGTATGAGAGAGA 1541  
QY 1511 GTATGAGAAATTCAGAGTGAATCAATGCTTGTGTCGAAGCCCAAGAAATTTCTGTA 1570  
DB 1542 ATATGAGAAATTTAAATCCCAATCAATGCACTTGTTCAAAGCCCTTAAATGTCTGA 1601  
QY 1571 AGAAGATGACATGCAAGATGGAACCCCTGAGCTGAGAAACATGTTGCTGTATCTC 1630  
DB 1602 AGAAGGTGAGTTATGCAAGTGAACAGCTGAGCTGAGAAATTAATACAGGAGACATCC 1661  
QY 1631 TGAATATGAGTCTTCTTCTGCGCAAGCGAGGCTTGACTGTGAGGAGAAATGAAT 1690  
DB 1662 AGAATATGACAGGCTTCTTGAAGCAAAATGTGTGAATGATGAGAGGCAATGAGCT 1721  
QY 1691 GCCAGATTTGTTATGTTTATAGAGAGAAACAGAGCTATTAACCTATTAAGAAAGC 1750  
DB 1722 CCGCGTTGTGATATGTTTCTGAGAAAGAGAGAGATTCAGAGCCCAAAAGGCG 1781  
QY 1751 TGTGCTATGATGATGATGCTCGAGTCTGTGATCAATGATGCTCATATTTGTT 1810  
DB 1782 TGTGCTATGATGATGATGCTGAGAGGTTTCAAGGTTCTTCAATGAGACCTTTCATCT 1841  
QY 1811 AAATGATTTGATCACTACATCAACAGAGAGCTTAAGAGAGAGAGATGTTT 1870  
DB 1842 GATCTGATTTGATCACTATTAATTAATTAACAGAGAGGCTTAAGAGAGAGATGTTT 1901  
QY 1871 TATGATGAGCCCTTACTAGAGAAAGGTTTGTATGATGATGATGATGATGATGAT 1930  
DB 1902 CCGATGAGACCAAACTCGGAGAGCAAGTTTGTATGTTGATGTTCCCAAAAGTTTGA 1961  
QY 1931 TGGATTTGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1990  
DB 1962 TGTATGATTAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2021  
QY 1991 GAAAGTTTGAATGATTCAGAGTCCAAATTAATGTTGATGATGATGATGATGATGAT 2050  
DB 2022 GAGAGTTTAAATGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2081  
QY 2051 GAGGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2110  
DB 2082 AACAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2120  
QY 2111 CTGCTGCGCCAGGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2170  
DB 2121 -----ACACAG 2174

QY 2171 TACCAACCCCAAAAGAGAAAGAAATTTATTTTTCAGAAAGAAAGAACCAATC 2230  
 DB 2175 TTCAAAGCTAAGAAAGATCGAGCAAAAAGAAATAGCAGCACTACCTCAACTGT 2234  
 QY 2231 CCTGATATGCTCTTGGTGAATTCAGCAAGCTGCTCCAGAGC-----TGAGAAATGA 2284  
 DB 2235 TCTGTATTCACCTTGATGACATGAAAGGAGGTGAGGTGCTGCTGTTTGAATGA 2294  
 QY 2285 AAAGCCGGATTTGTAATCAACAAAATTTAGAAAAGAAATTTGGCAACTCTTCTGTTT 2344  
 DB 2295 AAAGGCGCTTAATGTCCAAATGAGCGCTGAGAAAGCAATTTGACAGTCTGCTTTT 2354  
 QY 2345 TGTATCATCAACTTCTCGAAGATGTGGAACCTTGAAGAGTCAAGTCTGCTTCT 2404  
 DB 2355 TGTGCTCTACCTTATGGAATGGAATGTGTCTCTCTTCCAGCACTCCAGAAAAT 2414  
 QY 2405 TTTGAAAGAGCATATCATGTCTATGTTGTTGTTTGAAGCAACAAGACTGGGGAAA 2464  
 DB 2415 TCTCAAGAGCTATCATGTCTATGTTGTTGTTTGAAGATGATGATGATGAGGAAAT 2474  
 QY 2465 AGAGATTTGCTGATCTATGATCAGTACAGAAATATTTCACTGTTTCAAGATGA 2524  
 DB 2475 GAGATTTGATGATCTATGATCTGATGACAGAAATATTTGACTGGGTTCAAAATGA 2534  
 QY 2525 TTGTATGTTGGCGGTCAATTTACTGATACCTAAACGGGTTGACTTCAAGGTTCTGC 2584  
 DB 2535 TGCCCTGATGCGATGATCAATTTACTGATGCTTAAGCTTCCAGCTTCAAGGTTCTGC 2594  
 QY 2585 ACCCTGAAATCTTTGATGATGCTTCCAGAGGCTGCTGGGGCTCTGGGCTATGA 2644  
 DB 2595 TCTATCAATCTTTGATGATGCTTGAACCAAGGCTGAGGTTGAGTTGAGTTCACTGA 2654  
 QY 2645 GATCTCTTCAAGCAATCATTTGCCCTTTGATGAGGATGATGAGGCTGCAAAATTTT 2704  
 DB 2655 GATCTCTTCAAGGCAATTTGCTTATGATGATGATGATGATGATGATGATGATGATGAT 2714  
 QY 2705 GGAAGATTTTCTCATCACTCACTCGATGATGATGATGATGATGATGATGATGATGAT 2764  
 DB 2715 TGAGAGTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2774  
 QY 2765 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2824  
 DB 2775 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2834  
 QY 2825 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2884  
 DB 2835 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2894  
 QY 2885 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2944  
 DB 2895 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2954  
 QY 2945 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3004  
 DB 2955 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3014  
 QY 3005 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3061  
 DB 3015 CGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3074  
 QY 3062 AGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3121  
 DB 3075 TGAGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3134  
 QY 3122 GAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3181  
 DB 3135 AAATGAT 3194  
 QY 3182 GGGCCCTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3241  
 DB 3195 GGGACCACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3254

QY 3242 TCTCAAGTTTGGTGAAGCAAAACAGGACCAACGATGATGATGATGATGATGATGATGAT 3301  
 DB 3255 CTTCAAGGTTTGGTGAAGCAAAACAGGACCAACGATGATGATGATGATGATGATGATGATGAT 3314  
 QY 3302 TCTGCTGCTTCAATCTTCTGCTCTTGGGTTGAGTTGATGATGATGATGATGATGATGATGAT 3361  
 DB 3315 TCTGCTGCTTCAATCTTCTGCTCTTGGGTTGAGTTGATGATGATGATGATGATGATGATGAT 3374  
 QY 3362 TGAGTCTGCTTCTTGAAGAGTGTGTTGATGATGATGATGATGATGATGATGATGATGAT 3397  
 DB 3375 CACTGCGCGGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3410

## RESULT 5

US-08-960-048-1

Sequence 1, Application us/08960048C

Patent No. 6271443

GENERAL INFORMATION:

Applicant: Stalker, D. et al.

TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

TITLE OF INVENTION: Sequences

FILE REFERENCE: 15621/01/US

CURRENT APPLICATION NUMBER: US/08/960,048C

CURRENT FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/029,987

PRIOR FILING DATE: 1996-10-29

NUMBER OF SEQ ID NOS: 12

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic oligonucleotide

US-08-960-048-1

Query Match 28.8%; Score 1134.8; DB 3; Length 3328;  
 Best Local Similarity 66.1%; Pred. No. 5.8e-312;  
 Matches 1731; Conservative 0; Mismatches 827; Indels 60; Gaps 4;  
 QY 768 GCATATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 827  
 DB 488 GCTGAAGACATGGAATTCATTTGGAAGACAGGGTGAAGAGTTGGAAGAAAGAG 547  
 QY 828 GAGAGATGACCAAGACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 887  
 DB 548 AACAGAGAGAGAACCTGCAACACTAAGTTGAAGAGAGGCTGAAGATCCACTAG 607  
 QY 888 CCACTAATGA-----TGAGCAAGACAACTGTCCAGAAATTTCA 932  
 DB 608 CAACAAATGGAAGATMAACCGGACCGGATGCTTCCAGCCCTCTCACTAATAATCCA 667  
 QY 933 CTTTCATCAAGCAGATTAATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 992  
 DB 668 ATCCGAAAGACAGCTTGCACATACCGAACCGTATGATGATGATGATGATGATGATGATGAT 727  
 QY 993 GGGTCTTCTTCACTACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1052  
 DB 728 GGTCTTCTTCTTCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787  
 QY 1053 ATATCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112  
 DB 788 ACTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847  
 QY 1113 TGGTCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1172  
 DB 848 TGTATCTCTTGAAGAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907  
 QY 1172 GGGCCAGCATCTCACTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1232  
 DB 908 GGTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967  
 QY 1233 GAACCTCTTGGTCAACAAATGCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 1292



Db 968 GAGCCTCATGATTAAGCAAACTGCTGCTTCCATCTGCTGACCTGACCTACCCGGTG 1027  
Oy 1293 GATAAGGTTCTGTGATGTTTCTGATGATGCTGCAATGCTPACGTTGAAGATTA 1352  
Db 1028 GATAAGGCTCTGTTATATATGATGATGCTGGCCATGCTGACATTTGAATCTCTA 1087  
Oy 1353 TCTGAAATCTGAAATTTGCAAGAAATGGGTTCTTTCTGCAAAAGGTACAAATTTGA 1412  
Db 1088 GTAGAAACAGCCGACTTTGCAAGAAAGTGGTTCCATCTGCAAAAATTTTCCATTGA 1147  
Oy 1413 CCTGCGCTGCAAGATGATCTTCCAAAGAAATGATCTACTTGAAGACAAAGGTGCA 1472  
Db 1148 CCGCGGACCTGAGTTTACTTCTCAAGAAATGATTTACTTAAAGATTAAGTGCAG 1207  
Oy 1473 GCAAACTTTGTTAGGAGAGAGACATGAAAGAGATGAGAAATTCAGGTGGA 1532  
Db 1208 CCTCTTTTGTAAAGAAAGTGAAGCTATGAAAGATTTGAAAGATGACAAATTTGA 1267  
Oy 1533 ATCAATGCTTAGTTGCCAAAGCCGAAAGTTCTGAAAGAGATGACAAATGCAAGAT 1592  
Db 1268 ATCAATGCTTAGTTGCCAAAGGCTCAGAAACACCTGATGAAGATGACAAATGCAAGAT 1327  
Oy 1593 GGAACCCCTGCGCTGAAACCAATGTTGCTGATCATCTGCAATGATTCAGGTTCTCT 1652  
Db 1328 GGAATCTTGTGGCCAGAAATPACCCGGGTATCACTCTGCAATGATTCAGGTTTCTCT 1387  
Oy 1653 GGCACAAAGCGAGCCTTGACTGTGAGGAAATGAATGCGACGATGCTTATGTTTCT 1712  
Db 1388 GGATATGATGCTGCTGATGACATCGAAGAAATGAATTCCTCGACTGTTTACGTTCT 1447  
Oy 1713 AGAGAAACCAACAGGCTATPACATCATAGAAAGTGTGCTATGAATGCAATGCTC 1772  
Db 1448 AGAGAAAGAAACCTGCTCACCAACCAAAAAGGCTGCTGAAATGCTTGTGTT 1507  
Oy 1773 CGAGTCTCTGCTACTAACAATGCTCCATTTGTTAACTTGATTTGATGATCACTAC 1832  
Db 1508 AGGCTGTGCAATCTTACAAATGCTCCCTTCAATCTGATTTGATGACACATAT 1567  
Oy 1833 ATCAACAAACAGAGCTATPAAAGAAAGCAATGTTTATGATGAGCCCTTACTAGA 1892  
Db 1568 GTTAAACATPAGCAAGGAGTTAGGAGCAATGCTCTTCTGATGACCCCAAGTTGCT 1627  
Oy 1893 AAGAAAGTTTCTATGATGATGCTCCCTCAAAAGATTTGATGAGATGCAATGCA 1952  
Db 1628 CGAGATGATGCTATGCTAGTTCTCTCAAAAGATTTGATGAGATGAGATGATGA 1687  
Oy 1953 TATGCTAACCGAATGTTGCTTTTGTATCAATGAAAGTTGATGATGATGATGATG 2012  
Db 1688 TATGCTAACCGAATGTTTCTTTTGTATGATGAAAGTTGATGATGATGATGATG 1747  
Oy 2013 GATCCAAATTTATGTTGATGATGATGATGATTTAGAAAGGCAAGCTTATGATGAT 2072  
Db 1748 GGGCCAGTTTATGTTGGAACAGGTTGTGTTTCAATGAGCAAGCACTTATGATGAT 1807  
Oy 2073 GCGCCCAAAACAAAGAACCCATCAAGACTTGAAGCTGCTGCGCCCAAGTGTCTT 2132  
Db 1808 CCACCTTCAATGCCAA-----GTTTTCCCAAGTCAATCTCC 1843  
Oy 2133 TGCTGTGCTGTTTGGCAATAGAAAGAAAGAAAGATACCAAAACCAAAACAGAGAG 2192  
Db 1844 TCATCTCTGCTGTGTTGCTCCCGGGCAAGAAAGCACTTAAAGATCC----- 1890  
Oy 2193 AAAAAGTATTTATTTTCAAGAAAGAAAGAAACCAATCCCTGCAATGCTTGTGTA 2252  
Db 1891 -----ATCAGAGCTTATAGGAGTCAAAAGGGAAGAACTTGAATGCTGCAATCTTAA 1945  
Oy 2253 ATTGACGAAGCTGCTCAGAGAGCTGAAATGAAAGGCGGATTTGTTAAATCAACAAA 2312  
Db 1946 CTTAGGGAATATGCAATATGATGATGATGAAAGATCAATGTTATCTCTCAACCAAG 2005  
Oy 2313 TTAGAAAAGAAATTTGGCAATCTTCTGTTTGTATATCAGACTCTGAGAAATGAT 2372

Db 2006 TTTGAGAAAATTTTGGCTTATCTTCACTGCTTCAATTAATCACTAATGAGAAATGGA 2065  
Oy 2373 GGAACCTTGAAGAGTGAAGTCCGCTTCTCTTTTGAAGAAAGTATACATGATATAGT 2432  
Db 2066 GGAATGCTGATCTGCAACCTTCCACCTAATCAAGAAAGAAATTCATGATCAGC 2125  
Oy 2433 TGTGTTATGAGAACAGACAGACTGGGAAAAAGATTGGCTGATCTATGATGATGAT 2492  
Db 2126 TGTGCTATGAGAAAGAAAGATGATGAGGAGAAAGATGATGATGATGATGATGATG 2185  
Oy 2493 ACAGAAATATCTTACTGCTTTCAGATGATGATGATGATGATGATGATGATGATGATG 2552  
Db 2186 ACTGAGATATCTTAAACCGGCTCAAAATGCACTGCGAGGATGAGATGATGATGATG 2245  
Oy 2553 ATACTTAAACGGGTTGATTCAAAGGTTTCTGCACTCTGATCTTCAATGATGATGAT 2612  
Db 2246 ATGCCCTTAAAGCCAGATTCAAAGATGAGATGAGATGAGATGAGATGAGATGAG 2305  
Oy 2613 CAGGTGCTGAGTGGGCTTGGGCTTATGAGATCTTCTTCAAGATGATGATGATGATG 2672  
Db 2306 CAGGTTTCTGATGAGGCTCTTGGATCTGTTGAATTTTCTTAAAGAGAGATGATG 2365  
Oy 2673 TGTATGAGAT--GATGGGCTGTAATTTTGAAGAAATTTCTTCAATGATGATGAT 2729  
Db 2366 TGTATGAGCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2425  
Oy 2730 ATGCTGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2789  
Db 2426 ATGCTGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2485  
Oy 2790 TTATGACAGGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2849  
Db 2486 CTTCACAGGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2545  
Oy 2850 TCATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2909  
Db 2546 GGCCTTTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2605  
Oy 2910 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2969  
Db 2606 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2665  
Oy 2970 TTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3029  
Db 2666 TTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2725  
Oy 3030 ACATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3089  
Db 2726 ACTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2785  
Oy 3090 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3149  
Db 2786 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2845  
Oy 3150 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3209  
Db 2846 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2905  
Oy 3210 GCATTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3269  
Db 2906 TCTCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2965  
Oy 3270 AGACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3329  
Db 2966 AGACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3025  
Oy 3330 TGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3367  
Db 3026 TGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3063

RESULT 6  
US-09-838-586-1

```

; Sequence 1, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE OF INVENTION: Sequences
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-838-586-1

```

```

Query Match      28.8%; Score 1134.8; DB 4; Length 3328;
Beet Local Similarity 66.1%; Pred. No. 5.8e-312;
Matches 1731; Conservative 0; Mismatches 827; Indels 60; Gaps 4;

```

```

QY 766 GCATATGGGTATGGTATGTTGCTTGGAGGAAACGATGATGGAATTTGGAAGACAGACAA 827
DB 488 GCTTGAAGACATATGGAAATTCGATTGGAAAGAACAGGTGGAAGTTGGAAGAAAGAAAG 547
QY 828 GAGAGATGACACAGACGGGGAATGATGTGTGTGATGATGATGATGATGATGATGATGAT 887
DB 548 AACAGAAAGAGAGAGCTGCAACAACTAAGTTGAAAGAGAGCTGCAATCCACCTGAG 607
QY 888 CCACTATATGA-----TGAAGCAAGCAACAACTGTCGAGAAATTTCCA 932
DB 608 CAACAATATGAAATGAAACCGGACCGGATGCTCCAGCCCTCGACTAATATTTCCA 667
QY 933 CTTTCATCAAGCCAGATTAATCATATAGATGATTAATCATTAATGCGCTGAGTTTG 992
DB 668 ATCCGAAAGAGAGACTTCAACCATCCGAAACCGTATCATTAATGCGATTAATCATTTCT 727
QY 993 GGGTCTCTTCCACTACCGAGTATGATCCCGGTGAATGATGATGATGATGATGATGATGAT 1052
DB 728 GGTCTTTTCTTCATTAATGAGTAAACAAACCCGTTGACAGTCTTTGAGACTGTGCTC 787
QY 1053 ATATCTGTATCTGTGAATCTGTGTTGCCATGCTTGTGATCTTGTATCAATTTCCAAAG 1112
DB 788 ACTTCAGTATATGTGAATCTGTGTTGCAATTTTCTGGGTGTGATGATGATGATGATGATG 847
QY 1113 TGGTTCCTATGAGAGAGAGACTTACAGACCGGCTGTCACTGAGTTGCAAGAGAA 1172
DB 848 TGGTATCTGTTTACGAGGAACATATGACAGACTATCTGCAAGATATGAAAGAGAA 907
QY 1173 GGCACGCAATCTCACTGCTCCAAATTTGATTTCTTGTGATGATGATGATGATGATGATGAT 1232
DB 908 GGTGAACCTGATGAACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAA 967
QY 1233 GAACTCTTGTGTCACAACTGCTTCTATCTATCTATCTTTCGTTGATTTTCTGTT 1292
DB 968 GAGCCTCAATGATTTCTGCAATACCTGCTTTCATCTTTCGCTTTCGATACCGGAG 1027
QY 1293 GATTAAGTTTCTGCTATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1352
DB 1028 GATTAAGTTTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
QY 1353 TGTGAACAATCTGAATTTGCAAGAAATGAGTTTCTTTCGCAACGGTATCAATATTTGAA 1412
DB 1088 GTTGAACAAGCCACTTTTGAAGAAATGAGTTTCTTTCGCAACGGTATCAATATTTTGA 1147
QY 1413 CCGCGGCTTCAAGTGTATCTTCAACAGAAATGATGATGATGATGATGATGATGATGATGAT 1472

```

```

DB 1148 CCGCGGCACTGAGTTTACTTCTCAAGAGATTTACTTTGAAAGATTAAGTGCAG 1207
QY 1473 GCMAACTTTGTTAGGAGAGAGCAATGAAAGAGATGATGAAATTTCAAGGTGAG 1532
DB 1208 CCTCTTTTGAAGAAAGCTATGAAAGAGATTTATGAAAGATCAAAATTTGGA 1267
QY 1533 ATCAATGCTTATGTTCCAAAGCCAGAAATTTCTGAAAGAGATGAAATGCAAGAT 1592
DB 1268 ATCAATGCTTATGTTCCAAAGCTCAAGAAACCTTGAATGAAAGATGCAATGCAAGAT 1327
QY 1593 GGAACCCCTGAGCTGAAACAAATTTGATGATCATCTGGAATGATCAGGTCTTCCT 1652
DB 1328 GGAATCTTGTGCAAGAAATTAACCGGTGATCACTGCAATGATTCAGTTCCTT 1387
QY 1653 GGCMAAGCAGAGGCTTGAAGTGAAGAAATGAACCTGCAATGATGATGATGATGATGAT 1712
DB 1388 GGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1447
QY 1713 AGAGGAAAGACCAAGCTATTAACATTAAGAAAGCTGCTGATGATGATGATGATGATGAT 1772
DB 1448 AGAGGAAAGACCAAGCTATTAACATTAAGAAAGCTGCTGATGATGATGATGATGATGAT 1507
QY 1773 CGAGTCTGCTGATTAATCAAAATGCTCATATTTGTTAACTTGATGATGATGATGATGATGAT 1832
DB 1508 AGGTGCTGCAATCTTAACAAATGCTCTTCAATCTGCAATCTGATGATGATGATGATGAT 1567
QY 1833 ATCAACAAACAGCAAGCTATTAAGAAAGCAATGATGATGATGATGATGATGATGATGAT 1892
DB 1568 GTTAACAAATGCAAGCTATTAAGAAAGCAATGATGATGATGATGATGATGATGATGATGAT 1627
QY 1893 AAGAAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1952
DB 1628 CGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
QY 1953 TATGCTAACCGGAATGTTGCTTTTGTATGATTAATCAATGAAAGTTGATGATGATGATGAT 2012
DB 1688 TATGCTAACCGGAATGTTGCTTTTGTATGATTAATCAATGAAAGTTGATGATGATGATGAT 1747
QY 2013 GGTCCAAATTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2072
DB 1748 GGGCCAGTTTATGTTGGAAGAGTTGTTTCAATAGGCAAGCACTTATAGGCTATGAT 1807
QY 2073 GCGCCCAAAACAAAGAGCCACATCAAGCACTTGCATGCTGCGCCCAAGTGTCTTT 2132
DB 1808 CCACCTCAATGCCAA-----GTTTCCCAAGCATCCTCC 1843
QY 2133 TGTGTTGCTGTTGCAATAGAGAGCAAAAGAACTTACCAACCAAAACAGAGAG 2192
DB 1844 TCAATTTGCTGTTGCTGCTGCGGCAAGAAAGAACTTAAAGATCC----- 1890
QY 2193 AAAAAGTTATTTTTCAGAAAGAGAAAGCAATCCCTGATGATGATGATGATGATGATGAT 2252
DB 1891 -----ATCAGAGCTTATAGGATGCAAAAGGAAAGAACTTATGATGCTCATCTTTAAC 1945
QY 2253 ATTAAGCAAGCTGCTCAAGAGCTGAGATGAAAGGCGGATTTGTAATCAACAAAA 2312
DB 1946 CTAGAGAAATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2005
QY 2313 TTAGAAAGAAATTTGGCAATCTTCTGTTTGTAAATCAACCACTTCTGAGAAATGAT 2372
DB 2006 TTGGAAGAACTTTGGCTTATCTTCAAGCTTCAATTAATCAATGATGATGATGATGATGAT 2065
QY 2373 GGAACCTTGAAGAGTGAAGTCTGCTTCTTTCGAAAGAGATTAATGATGATGATGATGAT 2432
DB 2066 GGAAGTGGGCAATCTGCAACCTTCCACATTAATCAAGAAAGCAATTCATGATGATGATGAT 2125
QY 2433 TGTGTTATGAAGCAAGCAAGCAAGCTGAGAAAGATTTGCTGATCTATGATGATGATGAT 2492
DB 2126 TGTGCTATGAAGCAAGCAAGCTGAGGAGAAAGATTTGATGATGATGATGATGATGATGAT 2185
QY 2493 ACAGAAAGATTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2552
DB 2186 ACTGAGATATCTTAACCGGCTTCAAAATGCACTGCGAGATGAGATGATGATGATGATGAT 2245

```

QY 2553 ATACTAAGCGGTTCATTCAGAAAGTCTGCACTCGAATCTTTCAGATCGCTTCAC 2612  
DB 2246 ATGCCCTTAAGCCAGCAGCATTCAGAAAGATCTGCAACCATCATCTGTCTGATCGTTGAC 2305  
QY 2613 CAGGTCCTCGGTGGGCTCTTGGGCTCTATAGATCTCTTCAGCAATCATTTGCCCTT 2672  
DB 2306 CAGGTCCTCGGTGGGCTCTTGGGCTCTATAGATCTCTTCAGCAATCATTTGCCCTT 2365  
QY 2673 TGGATGCGGTAT--GGTGGCGGTGAAATTTTGGAAAGATTTTCTTACATCACTCC 2729  
DB 2366 TGGATGCGGTATGAGGTGTGTCTTAATAGCTTCAAGACTAGCATATATAAAGACC 2425  
QY 2730 ATCTGTATCTTGGACATCTATTCCTCTTGGCTTACTGTACATTTGCCCTCTGT 2789  
DB 2426 ATGTCTATCTTTCACATCTCTTCCATCTCATCTATGCTTCTTCTACACAGCATCTGT 2485  
QY 2790 TTATGACAGGGAATTTATCACTCCAGAGCTGAATTAATGTCAGGCTGGTTCATG 2849  
DB 2486 CTCTCAGAGGAAATTTATCTACCAAGCTCTCAAACTGGCAAGTGTCTCTTTCTT 2545  
QY 2850 TCACCTTTTATCTGCAATTTTGGCTACAGCATCTAGAAATGAGATGAGTGTGGA 2909  
DB 2546 GGCCTTTCTCTTCATTAATGCTGCTGCTTCTGAGCTCCAGTGGATGAGTGTGCA 2605  
QY 2910 ATTATGATCTGTGAGAGAAATGAGATCTTGGGCTTATGAGTGTGTCTTCACTC 2969  
DB 2606 ATTTGAGACTTATGCGATGAGAGATGAGTGTGGGCTATCGAGGCTTCAAGCCATCTC 2665  
QY 2970 TTTGCTGCTTCAGGAGCTTCTCAAGGCTATAGCTGAGTGTGATACAAAGCTTCA 3029  
DB 2666 TTTGCTGCTTCAGGAGCTTCTTAAAGATCTTGGGAGCTTGAACCACTTTTACTGTC 2725  
QY 3030 ACATCAAAAGGTGAGATGAGATGAGTCTCAGACTATATACATTCAAATGAGCTAC 3089  
DB 2726 ACTGCAAAAGCAGCTGATGAGATGAGTGTGGTGTGATGATGATGATGATGATGAT 2785  
QY 3090 TTATGATACCTCTTCAACCTTGTCTTATGAACTTCAATGAGTGTGTGTGTGTGT 3149  
DB 2786 CTCTCAATCTCTTCAACCACTCTCTCACTGCAATGATGATGATGATGATGATGAT 2845  
QY 3150 TCAATGAGATCAATGAGATGAGATGAGTGTGGGCTTGTGGGAGCTATCTT 3209  
DB 2846 TCCGATGCTTCAAAAGGATGAGATGAGTGTGGGAGCTTGTGGGAGCTTGTGTGT 2905  
QY 3210 GCATTTGGGTGATGCTCATCTTATCTCTTCAAGTGTGTGTGTGTGTGTGTGTGT 3269  
DB 2906 TCCCTCTGGGTGATCTCTCATCTTATCTCTTCAAGTGTGTGTGTGTGTGTGTGT 2965  
QY 3270 AGAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3329  
DB 2966 AGAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3025  
QY 3330 TGGGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3367  
DB 3026 TGGGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3063

RESULT 7  
US-09-221-013A-1  
; Sequence 1, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard B.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24

; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2248  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1887)  
US-09-221-013A-1

Query Match 22.5%; Score 884.6; DB 4; Length 2248;  
Best Local Similarity 68.5%; Pred. No. 6.2e-241;  
Matches 1277; Conservative 0; Mismatches 559; Indels 27; Gaps 3;

QY 1495 GAGCAATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1554  
DB 2 GAGCTATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 61  
QY 1555 CCCAGAAAGTCTGAAAGAGATGAGCAATGCAAGATGAAACCCCTGGCTGGAACA 1614  
DB 62 CACAGAAATCTCGAGAAAGCTGGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 121  
QY 1615 ATGTCTGATCATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674  
DB 122 ACCTTAAGATCATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181  
QY 1675 GTGAGGAAATGAACTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1734  
DB 182 CCGATGAAATGAACTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241  
QY 1735 ACCATCATGAAAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1794  
DB 242 AACACCAAAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301  
QY 1795 ATGCTCATATTTTGTAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854  
DB 302 ATGAGATATCTTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 361  
QY 1855 AGAAGCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1914  
DB 362 AAGAGCTATGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421  
QY 1915 TCCCTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1974  
DB 422 TCCCTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
QY 1975 TTTTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2034  
DB 482 TTTTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
QY 2035 GATGATATTTTAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2094  
DB 542 GTTGTGTTTAAATGAGGAGGCTCTATATGATGATGATGATGATGATGATGATGATGATGAT 601  
QY 2095 CATCAAGACTTGCATGCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2154  
DB 602 TAGAACCAATATTA-----TTGTCAAGAGCTGTGTGGGCTCA 640  
QY 2155 GGAAGCAAAAGAGACTACCAAAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2214  
DB 641 GGAAG 697  
QY 2215 AAG 2274  
DB 698 GAACTGATCTCAATGCTTCACTTTCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757  
QY 2275 CTGAGATGAG 2334  
DB 758 ATGATGATGAG 817

QY	2335	CTTCTGTTTTGTTACATCCACACTCTTCAGAAATGGGAAACCTGAAAGTGCATGTC	2394
Db	818	CGCGCGATTATTTATTCGGCCAACTTCATAGAAACAAGGGCGATTCCACCAACCAATC	877
QY	2395	CTGCTCTCTTTTGAAAGAACCTATACATGTCATTAGTTGGTTATGAGACAAAGACAG	2454
Db	878	CCGCTACTCTCTGAAAGAGCGCTATTCAATGTTATTAAGCTGTGGTTACGAAACAAGACTG	937
QY	2455	ACTGGGGAAAAGATTTGGCTGGAATCTATGGAATCAGTTACAGAAATATTTCTAACTGGTT	2514
Db	938	AATGGGGGAAAGAAATTTGGTTGGATCTATAGTTCCGTACGGAAGAAATATTTCTACTGGGT	997
QY	2515	TCAAGATGCAATGTCATGAGTTGGCGGTCAAATTTACTGCATACCTTAACGGGTTGCATCA	2574
Db	998	TCAAGATGSCAAGCCCGGGGTTGGATATGATCTACTGCAATCTCCACAGCCCTGGCTTCA	1055
QY	2575	AAGGTTGCACTCTGTAATCTTTTCAGATCGCTCTTCCACAGGTGCTTGGTGGGCTCTTG	2634
Db	1058	AGGGATTCGACCAATCAATCTTTCTGTATCGTTTGAACCAAGTTCTTGATGGGCTTTGG	1117
QY	2635	GGTCTATTGAATCTCTTTCAGCAATCATTTGCCCTCTTGGTATGGGATATGGTGGCGGTC	2694
Db	1118	GATCTATGAAATTTCTTTTAGCAGACATTTGTCTATCTGGTATGGTTTACATGGAAGGT	1177
QY	2695	TGAATTTTTGAAAGATTTTTCATACATCAATCCATCGTATTCCTTGAACATCTATTC	2754
Db	1178	TGAGACTTTTGAGAGGATCGCTTATATCAACACCATCGTCTATCTTATCATCTATCC	1237
QY	2755	CCCTCTGGCTTACCTGTAATCTTGCCTGCATCTGTTATTTGACAGGAAATTTATCACTC	2814
Db	1238	CTCTTATTTGCGTATTTGTAATCTTCCCGCTTTTGTCTCATCACCGACGATTCATCAATAC	1297
QY	2815	CAGAGCTGAATATATGTTCCAGCGCTGTGGTTCATATGCACTTTTATCTGCATTTTGGCTA	2874
Db	1298	CCGAGATAGCAATCACCGCAGTATTTGGTTATCTTACTCTTCAATCTCAATTCGTGTGA	1355
QY	2875	CGAGCATCTAGAAATGATGAGAGTGTGTGAAATTGATGACTGTGTGAGAAATGAGC	2933
Db	1358	CTGGAATCTGAAACTGAAATGGAACGGGTGTGACGATGAGGATTTGGTGGAGAACCAAC	1411
QY	2935	AGTCTGGGTCATTGGAAGGTGTGTCTCAACCTCTTGGTGTGTGTCAGGGACCTTCA	2994
Db	1418	AGTCTGGGTCATTGGAAGGTGTGTCTCAACCTCTTGGTGTGTGTCAGGGACCTTCA	1477
QY	2995	AGGTCATAGCTGTGTGTATACAAAGCTTCAACCGTGACATCAAAAGGATG--GATGATG	3053
Db	1478	AGGTTCTTGCTGGTATCAACACCAACTTCAACCGTTACATCTTAAGCCACAACAAAATG	1533
QY	3052	AGGAGTCTCAGAGCTATATACATTCAAAATGAGACTAATGATGATGATCTCTACACT	3111
Db	1538	GGGATTTTGCAAAATCTTCAATCTTCAATGAGCAGCTTCTCAATCCACCAACACACCG	1597
QY	3112	TGCTTCTATTGAACCTCAATGGGTGTGTGCGCTGGGCTTCAAAATGCGATCAATTAAGGAT	3177
Db	1598	TCTTACTTGTAACTCATAGGCAATTTGGCTGGGTGTCTCTTATCTGTAAAGAGTGGCT	1655
QY	3172	ATGAGTCATGGGGCCCCCTCTTGGGAAGCTATTCTTTGCATTTTGGTGAATTTGCATC	3233
Db	1658	ACCAAGTCGTGGGGTCGCGCTTTTCGGGAAGCTCTTCTTGCGCTTATGGGTTATGGCCATC	1711
QY	3232	TTTATCCCTTTCTCAAAGTTTGGTTGGAAAGCAAAACAGGACACCAACGATTTGTCATG	3297
Db	1718	TCTACCCCTTTCTTGAAGGTCTGTTGGAAAGCAAAACCGAACAACCAACATCGTCATG	1777
QY	3292	TCTGGTCATTTCTGTGGCTTCAATCTTCTCGCTCTTGGGTTGGATTTGATCTTTCC	3353
Db	1778	TCTGGTCTGTTCTTCTCGCTCTCAATCTTCTGTCTTGGGTCAGATCAATCCCTTTG	1833
QY	3352	TTTG 3354	
Db	1838	TGG 1840	



Db 5381 ATACGTAAGAAACATGTTACATCGTTTGTTCATTTTCAGGCTTTTACGGCAT 5440  
Qy 1659 AGCGAGGCGCTTGA CTGTGAGGAAATGAATGCCAGATTTGTTATGTTTCTAGAG 1718  
Db 5441 AGTGGGGGTCTGATATACCATGAGAAATGAGCTGCTAGACTCATCTATGTTCTCGTAA 5500  
Qy 1719 AAAGACACAGGCTATTAACCATCATTAAGAAACGTGTGTATGATGCACTTGG----- 1770  
Db 5501 AAGGCGCTGTGATTTCAACACCAAAAAGGCTGAGCTATGATGCAATGGTTTGTAA 5560  
Qy 1771 ----- 1770  
Db 5561 CTTTCAGAACTCTATGTGTCTCTATTTATTTCTCTTGTTCACCTGCTTAAGAAACGTT 5620  
Qy 1771 ----- 1770  
Db 5621 TTCTGTGTAGCCGTGTGCTTACATTTCTTTTCTTCTAGCTATGTGTCTCTCTAAT 5680  
Qy 1771 ----- 1770  
Db 5681 TAGTATCTTTTACTTTGACAGATCCGTGTATCTGTGTCTTCTTACCAATGAGCATATCT 5740  
Qy 1808 GTTAACTTGATTTGTATCTACTCATCAACAAGAGGCTATTAAGAGCAATGTG 1867  
Db 5741 TTGAACTGTGATTTGTATCTACTTTTATTAACAAGTAAAGGCTATGAGTGTG 5800  
Qy 1868 TTTTATGATGACCTTTTACTAGAGAAAGGTTGTATGTATGACAGTTCCCTCAAAAT 1927  
Db 5801 TTTTATGATGACCTGCTATTTGAAAGAGTGTGTATGTCCAGTTCCCTCAACGTTT 5860  
Qy 1928 TGATGGGATTTATGAGCCATGACCGATATGCTAACCGGAATGTGTCTTTTGTAT 1982  
Db 5861 TGACGGTATTTATTTGACGATGATATGATGCAAGAAATATGATTTTTCATGTGAG 5920  
Qy 1983 ----- 1982  
Db 5921 TATCACTTCCCATTTGCTTTTGTCTTTTGTCTTATTTTGTGATTTTACCTCT 5980  
Qy 1983 ----- 1982  
Db 5981 TTCTGCTATGCGCTGATGATATTGTTCTCTTGGCAGATTAAATGAAAGGCTGTG 6040  
Qy 2002 ATGCTATTCAGGCTTCAATTTATGTTGTACTGTATGTTTAAAGAGGAGCATAT 2061  
Db 6041 ATGCTATTCAGGCTTCAATTTATGTTGTACTGTATGTTTAAAGAGGAGCTAT 6100  
Qy 2062 ATGCTATTCAGGCTTCAATTTATGTTGTACTGTATGTTTAAAGAGGAGCTAT 6100  
Db 6101 ATGCTATTCAGGCTTCAATTTATGTTGTACTGTATGTTTAAAGAGGAGCTAT 6100  
Qy 2116 GAGCCAGTGTGCTTTTGTCTGTGTCTGCTTTGCAATGAGAACAAAGAGCTA 2175  
Db 6161 GTTCCGGGTCAAGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 6220  
Qy 2176 AACCCAAACAGAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2227  
Db 6221 GCATCAACAGAGTGTCTCAATGCTCCATTTTCAATATGAGAGCATCATAGAGGTT 6280  
Qy 2228 ----- 2228  
Db 6281 TTGAAGGTTTATGAGCTATGTTATTAATCACTCTTTTATGTAATGATTTATGT 6340  
Qy 2244 CTTGTGAAATTTGACGA-----GCTGCTCCAGAGCTGAGATGAAAGGCGGTATG 2298  
Db 6341 GATGCTGAAATCTTAATCTTGTATGAGAGTATGATGAGAGAGGCTATTTCTAA 6400  
Qy 2299 TAAATCAACAAATTTGAAAGAAATTTGCGCAATCTTCTGTTTGTATCATCAAC 2358  
Db 6401 TGTCCCAAGAGAGTGTAGAGAGGTTTGTGTGATGCTCCGCTATTTATGCGCAACCT 6460  
Qy 2359 TTCTGAGATGTGTGAACTTTGAAGAGTGTGCAAGCTTCTCTTTTGAAGAGCTA 2418

Db 6461 TCATGAAACAGGCGGATTCACCAACCAATCCCGTACTCTTTCGAAGAGCTA 6520  
Qy 2419 TACATGCTATTTGTTGTATTAAGAACAGACAGACTGGGAGAAAG----- 2466  
Db 6521 TTGATGTTAACTGTGTTTGAAGAACAGACTGAATGGGCAAAAGAGTCAATTTTC 6580  
Qy 2467 ----- 2466  
Db 6581 AAATGACCTACGAATCTTCTTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6640  
Qy 2467 ----- 2467  
Db 6641 GGCATTTTGTATGATTTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 6700  
Qy 2515 TCAAGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2574  
Db 6701 TCAAGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 6760  
Qy 2575 AAGGTTCTGACCTGGAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2634  
Db 6761 AAGGATCTGACCAATCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6820  
Qy 2635 GATCTATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2694  
Db 6821 GATCTATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6880  
Qy 2695 TGAATTTTGAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2754  
Db 6881 TGAATTTTGAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6940  
Qy 2755 CCTCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2814  
Db 6941 CCTCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 7000  
Qy 2815 CAGAGCTGAATATGTT----- 2831  
Db 7001 CCGAGGTTTGAATCTGACCAACATCTTATTTTGAATCCCATTTTGTGAATGC 7060  
Qy 2832 ----- 2832  
Db 7061 ATTTTGTGCTATCATCATTTGTTGAGATTAAGCACTAGCGGATTTTGTGCTATTTCTA 7120  
Qy 2853 CTTTATTCGATTTTGTGCTATGAGCATTTCTTGAATGAGATGAGTGTGTGAAT 2912  
Db 7121 CTTTATTCGATTTTGTGCTATGAGCATTTCTTGAATGAGATGAGTGTGTGAAT 7180  
Qy 2913 GATGATGAGTGTGAGATGAGCTTCTGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2972  
Db 7181 GAGGATTTGAGGAGAGAGAGAGCTTCTGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7240  
Qy 2973 GCTGTGCTCAGGAGCTTCTCAGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3032  
Db 7241 GCTGTGCTCAGGAGCTTCTCAGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7300  
Qy 3033 TCAAGG---GTGAGATGATGAGAGTCTGAGCTATTAATCAATGAGTCAAC 3089  
Db 7301 TCTAAAGCCACAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7360  
Qy 3090 TTAATGATCTCTTCTCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3149  
Db 7361 CTTCTATTCACCAACACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 7420  
Qy 3150 TCAATGCTATCAATACAGATTAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3209  
Db 7421 TCTTATGCTGTAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7480  
Qy 3210 GCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3269  
Db 7481 GCCTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7540  
Qy 3270 AGGACACCAAGATTTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGAT 3329  
Db 7541 CGAACACCAACATGCTATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGAT 7600

```

Oy      3330  TGGGTTGGATGATCCCTTCTGCTG 3354
          ||||| ||||| ||||| |||||
Db      7601  TGGGTGAGATCATCCCTTGTGG 7625

RESULT 10
US-09-221-013A-4
: Sequence 4, Application US/09221013A
: Patent No. 6495740
: GENERAL INFORMATION:
: APPLICANT: Arioli, Antonio
: APPLICANT: Williamson, Richard E.
: APPLICANT: Betzner, Andreas S.
: APPLICANT: Peng, Liangcai
: TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
: FILE REFERENCE: 96-98
: CURRENT APPLICATION NUMBER: US/09/221,013A
: CURRENT FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: PCT/NV97/00402
: PRIOR FILING DATE: 1997-06-24
: PRIOR APPLICATION NUMBER: AU P00699
: PRIOR FILING DATE: 1996-06-27
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 4
: LENGTH: 5009
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(5009)
: OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4

```









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 15, 2003, 13:00:59 ; Search time 27 Seconds

(without alignments)  
3868.118 Million cell updates/sec

Title: US-09-720-383c-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNRNELVYI.....DEFLAKDGPPLLEEGCLDN 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4278.5	72.3	1084	2 T08583	cellulose synthase
2	4209.5	71.1	1088	2 H84604	cellulose synthase
3	4202	71.0	1081	2 T52028	cellulose synthase
4	3967.5	67.0	1026	2 T51579	cellulose synthase
5	3945.5	66.7	1081	2 T05351	cellulose synthase
6	3875.5	65.5	1065	2 T52054	cellulose synthase
7	3782	63.9	1065	2 F84649	cellulose synthase
8	3469.5	58.6	974	2 T10797	cellulose synthase
9	3329.5	56.3	958	2 T04870	cellulose synthase
10	2907	49.1	685	2 T10800	cellulose synthase
11	2330.5	39.4	1181	2 D86157	hypothetical prote
12	2311	39.1	1145	2 T51546	cellulose synthase
13	2256	38.1	1111	2 T05646	hypothetical prote
14	2166.5	36.6	1036	2 D84741	probable cellulose
15	2118.5	35.8	583	2 T02209	cellulose synthase
16	2114	35.7	979	2 C86446	probable cellulose
17	1056.5	17.9	757	2 T02561	cellulose synthase
18	1046	17.7	712	2 T02552	cellulose synthase
19	1040	17.6	755	2 T02553	cellulose synthase
20	993.5	16.8	748	2 T02560	cellulose synthase
21	959.5	16.2	689	2 T08918	hypothetical prote
22	952	16.1	727	2 T08920	hypothetical prote
23	944.5	16.0	686	2 T08919	hypothetical prote
24	901	15.2	828	2 E71417	hypothetical prote
25	829	14.0	710	2 B71417	hypothetical prote
26	414.5	7.0	336	2 T08591	cellulose synthase
27	402.5	6.8	693	2 AF2275	cellulose synthase
28	393	6.6	322	2 T12093	cellulose synthase
29	359.5	6.1	322	2 T08592	cellulose synthase

30	256.5	4.3	759	2 D70422	cellulose synthase
31	252	4.3	754	2 A43735	beta protein - Ace
32	248	4.2	874	2 AB0985	probable polysacch
33	239.5	4.0	1596	2 T31338	cellulose synthase
34	239	4.0	888	2 E91180	probable cellulose
35	238	4.0	692	2 S47754	hypothetical prote
36	238	4.0	888	2 H65151	hypothetical 101.6
37	232	3.9	888	2 G86026	probable cellulose
38	227.5	3.8	729	2 A98320	cellulose synthase
39	227.5	3.8	729	2 AD2963	cellulose synthase
40	227.5	3.8	861	2 T39714	cellulose synthase
41	211	3.6	664	2 A95889	probable cellulose
42	201	3.4	723	2 C36963	cellulose synthase
43	194	3.3	768	2 B97083	glycosyltransferase
44	175	3.0	654	2 AB2932	beta 1,3 glucan sy
45	175	3.0	654	2 D98350	hypothetical prote

## ALIGNMENTS

## RESULT 1

T08583 cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana

N/Alternate names: protein T22F8.250

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 11-Jun-1999 #sequence \_revision 11-Jun-1999 #text \_change 22-Oct-1999

C/Accession: T08583; T09014

R/Bevan, M.; Zimmermann, W.; Gruenewald, A.; Mamut, R.; Bancroft, I.; Mewes, H.W.; M

submitted to the Protein Sequence Database, May 1999

A/Reference number: Z16442

A/Accession: T08583

A/Molecule type: DNA

A/Residues: 1-1084 <BEV>

A/Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.250

A/Experimental source: cultivar Columbia; BAC clone T22F8

R/Artoli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.;

Science 279, 717-720, 1998

A/Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.

A/Reference number: Z13745; MUID:98111412; PMID:9445479

A/Accession: T09014

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1084 <ART>

A/Cross-references: EMBL:AF027173; NID:g2827140; PIDN:AAC3935.1; PID:g2827141

A/Experimental source: cultivar Columbia

C/Accession: T08583

A/Map position: 4 156/3; 191/1; 254/2; 343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3;

A/Intons: 27/3; 156/3; 191/1; 254/2; 343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3;

C/Keywords: cell wall synthesis; glycosyltransferase; hexosyltransferase

Query Match 72.3%; Score 4278.5; DB 2; Length 1084;

Best Local Similarity 71.3%; Pred. No. 0;

Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;

QY	1	MEASAGLVAGSHNRNELVIRRDGPGRFPPEQNGQVQICGDVGLAPGDDPFACNE	60
DB	1	NMTGRLINGSHNRNEFLVINDSARIRSVGLSGQTQICGDELTVSSLEFACNE	60
QY	61	CAFPVCRDCYERYRRBGTGNCPOCKTRRYRLGRCRVTVG-DEEDGVDLIDNEFWMDGHD	119
DB	61	CAFPVCRPCYERYRRBGTGNCPOCKTRRYRLGRCRVTVG-DEEDGVDLIDNEFWMDGHD	119
QY	120	SGSVAESMLYGHMSYGRGDDPRGAQAQOLNPNVPLTNGQVNDLPPEQHALVPSFMCG	179
DB	120	PEHAALALSSRLNLTGRGLDSAPP-----GSQIPLLTYCDDADMTYSRRHLIVPSTG	174
QY	180	GSKRIHPLYPADPSLPVQPRSDPSKDLAAYGVGVAERMENMKORO-ERM---HOT	234
DB	175	YGNRYVPAPFTTSSAPPOARSMPQKDIIEYGVGVAMQDRHEWKRQGEGLQYIKHIG	234
QY	235	GNDGGG---DDGDDADPLMDENARQQLSRKILPSSQINPYRMIIIRLVVVGFFPHYR	290

```

Db      235 GNNGRSGNDDDELDDPMDMBGRQPLSRKLPTRSSRINDRYMLICRLATLGLFPHYR 294
Qy      291 VMHVPNDAPALMLISVICIWMFAMSWILDOFPKPFPIERETYLDRLSLRPDKEGQSOQA 350
      295 ILHVPNDAYGLMLSVICIMFAMSWILDOFPKPFPIERETYLDRLSLRPDKEGQSOQA 354
Qy      351 PIDEFVSTVDPKPEPLVTTNTVLSISVDYVDKVCYVSDGQAMLTPEALSSEFA 410
      355 PVDVFVSTVDPKPEPLVTTNTVLSISVDYVDKVCYVSDGQAMLTPEALSSEFA 414
Qy      411 KKWVPCKRYNIEERAPBEMVFOQKIDYLDKVAANFPERBAMREVEEPVRINALVAK 470
      415 KKWVPCKRYNIEERAPBEMVFOQKIDYLDKVAANFPERBAMREVEEPVRINALVAK 474
Qy      471 AOKVBEQMTMODSTPMPGNVNRDHPGNIQVFLQSGGLDCEGNELPRLVYVSREKPGY 530
      475 AOKVBEQMTMODSTPMPGNVNRDHPGNIQVFLHSGVROTDGNEPLRLVYVSREKPGY 534
Qy      531 NHHKAGAMNALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYQ 590
      535 DHHKAGAMNSLIIVSAVLSNAPYLNLDCDHYINNSKAIRESMCFMMDPQSGKKVCYQ 594
Qy      591 PPORFPGIDRHDRIYANRVVFPDINMKGLDGIQGPYVGTGCVFRQALVYDAPKTKKP 650
      595 PPORFPGIDRHDRIYANRVVFPDINMKGLDGIQGPYVGTGCVFRQALVYDAPKTKKP 654
Qy      651 PSRTCNCPKWCFCFCCCGFGRNKKOKTKTKTEKKKLLFFKKEENOSPAYALGEIDEA--A 708
      655 PGKTCNCPKWC-CLCC-GLRKKSKT--KAKDKKT--NTKETSQKHLEBNVDEGIYV 706
Qy      709 PGANEKAGIYNQOKLEKKFGQSSVFTSTLLENGGTLKASPSLSLKEAHHVSCGYED 768
      707 PVSNEKRSSEATOLKLEKKFGQSSVFTSTLLENGGTLKASPSLSLKEAHHVSCGYED 766
Qy      769 KTDGKEIGMIVGSTEEDILTFGMKHCHGRMSIYCIPIKRVAFKGSAPLNTSDRLHQLVLRW 828
      767 KTDGKEIGMIVGSTEEDILTFGMKHCHGRMSIYCIPIKRVAFKGSAPLNTSDRLHQLVLRW 826
Qy      829 ALGSEIEFFSNHCPMLWGYGGGLKFLERFSYINSIYVPMWTSIPPLAYCTLPALICLTGKF 888
      827 ALGSEIEFFSNHCPMLWGYGGGLKFLERFSYINSIYVPMWTSIPPLAYCTLPALICLTGKF 886
Qy      889 ITPELANVASLWMSLFCIPATSIILEKMSGVGIDMWNRQFVWIGVSSHLPVAFQ 948
      887 IVPESISYAGILFIMLMPFSIAVTGILEWQMGVGDIDMWNRQFVWIGVSSHLPVAFQ 946
Qy      949 LTKVIAGVDSFTYTSKGGDEBFSELYTFKMTLLIPPTLLLNFIIVAGVSNLNN 1008
      947 LTKVIAGVDSFTYTSKGGDEBFSELYTFKMTLLIPPTLLLNFIIVAGVSNLNN 1006
Qy      1009 GYESWGPFLFGKLFPAFWIVLHLYPFLKGLVGRONRTPPIVMSILLASIFSLMVAIDP 1068
      1007 GYESWGPFLFGKLFPAFWIVLHLYPFLKGLVGRONRTPPIVMSILLASIFSLMVAIDP 1066
Qy      1069 FLAYODGFLBEGCGLDC 1085
      1067 FLAYODGFLBEGCGLDC 1082

```

## RESULT 2

H84604  
 probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C/Accession: H84604  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Holtz, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;  
 Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: H84604

```

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1088 <STO>
A/Cross-References: GB:AE002093; NID:94417271; PIDN:AAD20396.1; GSPDB:GND01339
C/Genetics:
A/Gene: At2g21770
A/Map position: 2

Query Match      71.1%; Score 4209.5; DB 2; Length 1088;
Best Local Similarity 70.2%; Pred. No. 1.7e-314;
Matches 769; Conservative 127; Mismatches 179; Indels 21; Gaps 10;

Qy      1 MEASAGLVGSHNRNLVVRDQDGPKPRPQNGOVQICGDVDVGLARGPVYACNE 60
      1 MNTGRLIAGSHNRNLVLRNADTYARISAEELSQTCTICDELETDNGSPFIACNE 60
Qy      61 CAEPVCRDCEYERREBGTQNCPOCKTRYKALKCCQVYTGDEBDEGVDLDNENFMWGHDS 120
      61 CAEPVCRDCEYERREBGTQNCPOCKTRYKALKCCQVYTGDEBDEGVDLDNENFMWGHDS 118
Qy      121 QSVAESML-YGHMSYGRGDGPNAGPQAFQALNPVPLLTNGQVDDIIPREGHALVPFMSG 179
      119 EHYTEAALYYMRINTGRGTDEVSHTLSASPSGEVPLITYCEDSDMWSDRHALLVPPSTG 178
Qy      180 GGRHHPRLPYADPSLPVQPSMDSKDIAAYGVSVAKTERMENKQRO-ERN----HQT 234
      179 LGNRVHVHPFTDSFASIHTRPMVPOKDLTVYGVSAVMDRMVWKKQLEKLVQVKNR 238
Qy      235 GNDGCG-----DDGDDADPLAMDARQOLSRKIPLPSSQINPRMIIIRLVLAGPFFHY 289
      239 VNDGDDGGLFVDELDBGLPMDQEGRLPSRLKPISSRINPRMLIFGLALIGLGFHY 286
Qy      290 RVHVPNDAPALMLISVICIWMFAMSWILDOFPKPFPIERETYLDRLSLRPDKEGQSOQA 349
      299 RILHPVNDAPALMLISVICIWMFAMSWILDOFPKPFPIERETYLDRLSLRPDKEGQSOQA 358
Qy      350 APIDFVSTVDPKPEPLVTTNTVLSISVDYVDKVCYVSDGQAMLTPEALSSEFA 409
      359 APIDFVSTVDPKPEPLVTTNTVLSISVDYVDKVCYVSDGQAMLTPEALSSEFA 418
Qy      410 AKKVPCKRYNIEERAPBEMVFOQKIDYLDKVAANFPERBAMREVEEPVRINALVAK 469
      419 AKKVPCKRYNIEERAPBEMVFOQKIDYLDKVAANFPERBAMREVEEPVRINALVAK 478
Qy      470 KAOVBEQMTMODSTPMPGNVNRDHPGNIQVFLQSGGLDCEGNELPRLVYVSREKPG 529
      479 KAOVBEQMTMODSTPMPGNVNRDHPGNIQVFLQSGGLDCEGNELPRLVYVSREKPG 538
Qy      530 YNHKKAAGAMNALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPQSGKKVCY 589
      539 YNHKKAAGAMNSLIIVSAVLSNAPYLNLDCDHYINNSKAIKEAMCFMMDPQSGKKVCY 598
Qy      590 QPQRFPGIDRHDRIYANRVVFPDINMKGLDGIQGPYVGTGCVFRQALVYDAPKTKKP 649
      599 QPQRFPGIDRHDRIYANRVVFPDINMKGLDGIQGPYVGTGCVFRQALVYDAPKTKKP 658
Qy      650 PPSRTCNCPKWCFCFCCCGFGRNKKOKTKTKTEKKKLLFFKKEENOSPAYALGEIDEA 709
      659 PPSRTCNCPKWC-CLCC-GMRK-KTKGVYDQNRK---KPKETSQKHLEBNVDEGIYV 711
Qy      710 GANEKAGIYNQOKLEKKFGQSSVFTSTLLENGGTLKASPSLSLKEAHHVSCGYED 769
      712 VTNANNSSETAOLKLEKKFGQSSVFTSTLLENGGTLKASPSLSLKEAHHVSCGYED 771
Qy      770 TDMGKEIGMIVGSTEEDILTFGMKHCHGRMSIYCIPIKRVAFKGSAPLNTSDRLHQLVLRW 829
      772 TDMGKEIGMIVGSTEEDILTFGMKHCHGRMSIYCIPIKRVAFKGSAPLNTSDRLHQLVLRW 831
Qy      830 LGSIEIEFFSNHCPMLWGYGGGLKFLERFSYINSIYVPMWTSIPPLAYCTLPALICLTGKFI 889
      832 LGSIEIEFFSNHCPMLWGYGGGLKFLERFSYINSIYVPMWTSIPPLAYCTLPALICLTGKFI 891
Qy      890 TPELANVASLWMSLFCIPATSIILEKMSGVGIDMWNRQFVWIGVSSHLPVAFQGL 949

```

Db 892 VPEISNVAIGILFLMFMSIAVTGLIEMQKIGIDDMRNBOFWVIGVSSHLFALFQGL 951  
Qy 950 LKVIAGVDTSTFTVTKSGDDEEFSELYTKKTTLLIPPTLLLPFGVAVASNAING 1009  
Db 952 LKVLGAVSTNFTVTSKADDDGSESLYIFKMTSLIPPTLLIINIVGIVVSAIING 1011  
Qy 1010 YESMGPLFGKLEFPAFVIVHLVPLKGLVGRNRTPTIYVMSILLASIFSLMWRIDPF 1069  
Db 1012 YDSWOPFLGRFLPALMIVHLVPLKGLGKODRVPITILVMSILLASITLLMVRVDPF 1071  
Qy 1070 LAKDGPFLBECGLDC 1085  
Db 1072 VSK-DGPVLEICGLDC 1086

RESULT 3  
T52028  
cellulose synthase [imported] - Arabidopsis thaliana (fragment)  
C.Species: Arabidopsis thaliana (mouse-ear creas)  
C.Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C.Accession: T52028  
R.Joshi, C.  
submitted to the EMBL Data Library, May 1998  
A.Reference number: Z25890  
A.Accession: T52028  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: mRNA  
A.Residues: 1-1081 <JOS>  
A.Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match 71.0%; Score 4202; DB 2; Length 1081;  
Best local Similarity 70.8%; Pred. No. 6.5e-314;  
Matches 770; Conservative 126; Mismatches 172; Indels 20; Gaps 10;

Qy 7 LVAGSHNNELVIRDDPGPKPREQNGQVCOICGDDVGLAPGDPFVACNECAFVC 66  
Db 4 LIAGSHNNEFLIADENARIRISVQLSGTCCQCRDEILTVDGEFPAACNECAFVC 63  
Qy 67 RDCYEYERREGTONCPCKTRRYKRLKGCQVVTGDEEDGVDDLNEFMN--DGHDSQVA 124  
Db 64 RPYCYERREGNOACPOCKTRFKRLKGSFVGVGDEEDDIDLDNEFEYGNNGIGFDQVS 123  
Qy 125 ESMYLGHSYSG-RGDDPKAGAPAFOLNPNVPLITNGQVNDIPPEQHAL-VSPFMGGGSK 182  
Db 124 EGMISIRRNSGFPQSDLPASPPGSG---IPLLITGDEBDVEISSSRHALIYPPSGGGGN 179  
Qy 183 RIHPLPYADPSLPVQPRSMDEPKDLAAYGVGSVAMKEMMKORQ-PRMQHTGNDGGSD 241  
Db 180 RVHPVSLSDPTVAHRRLMVPQKDLAVYGVGSVAMKDRWEKRRKQNEKLQVVRHGGPD 239  
Qy 242 --DGDADLPLMDEARQOLSRKILPSSQINPYRMIIRLVVVGFFPHYRVMHFVNDAF 299  
Db 240 FEDGDADLPMMWDEKROPLSMKIPKSSKINPYRMLVILVITLGLFFHYRIHFVNDAY 299  
Qy 300 ALMLISVCEITFMAMSLDOPPKMFPITERETYLRLSLRKYEKKEKPSGLSPVDVSTV 359  
Db 300 ALMLISVCEITFMAMSLDOPPKMFPITERETYLRLSLRKYEKKEKPSGLSPVDVSTV 359  
Qy 360 DPLKEPLVTNTVTIISVDVVDKVSVCYSDDAAMLTFPALSSEBPAKKWVPCKR 419  
Db 360 DPLKEPLVTNTVTIISVDVVDKVSVCYSDDAAMLTFPALSSEBPAKKWVPCKR 419  
Qy 420 YNIBRAEWYFOQKIDYLDKQVAAVFYERRAMKREYEEFKVRIINALVAKAKVPEEGM 479  
Db 420 YCIBRAEWYFCHMDYLDKQVAAVFYERRAMKREYEEFKVRIINALVATQKVPEDGM 479  
Qy 480 TMQDGTTPWGNVNRDHPGMIQVFLQSGGLDCEGNEPLPLVIVSREKRGPIVNHKKAGAM 539  
Db 480 TMQDGTTPWGNVNRDHPGMIQVFLQSGGLDCEGNEPLPLVIVSREKRGPIVNHKKAGAM 539  
Qy 540 NALVVSIVLTMAVPLMLDGDHYTNNSKAIKEAMCFMMDPLGKKVVCVQPPGPFDDID 599  
Db 540 NALVVSIVLTMAVPLMLDGDHYTNNSKAIKEAMCFMMDPLGKKVVCVQPPGPFDDID 599

Db 540 NSLIRVSGVLSNAPYLLVDCDHYTNNSKALRBAMCFMMDPQSGKKICVQPPGPFDDID 599  
Qy 600 RHDRYARNRVFPDINMGGLDGIQPIYVGCYPPROALYGDAPTKKPPSPRTQNCWP 659  
Db 600 RHDRYSNNVVPFDINMGGLDGIQPIYVGCYPPROALYGDAPTKKPPSPRTQNCWP 659  
Qy 660 KMCFCCCFGNRKOKTKTPTEKKKLLFPKKEENSPAVYALGEIDEAA--PGAENEKAG 717  
Db 660 KMCL--LCFSGRRKRRKAKTVAAADKK---KOREASKIHLLENIEBGRGHVLTNEOST 713  
Qy 718 IVNQDLKEKFGQSSVPFTSTLNGTLLKASAPASLKEA1HYISCGYEDKTDWKEIG 777  
Db 714 EAMQMKLOKCKYGSPVFVVASARLENGGWARASAPCLLKBA1QVYSREYEDKTEGKEIG 773  
Qy 778 WYGSYVEDILITGFRMHGHRSTYICPKRAVFGSAPLANSDRLHOVLRNALSGTIEFP 837  
Db 774 WYGSYVEDILITGFRMHSRHRVYCTPKLAFGSAPINISDRHLQVLRNALSGVETFL 833  
Qy 838 SNHCPMTWYGGGKFLERFYSIYVPMTSIPILAYCTLPAICLTGKRTPELANNVA 897  
Db 834 SRHCPMTWYGGGKFLERFYSIYVPMTSIPILAYCTLPAICLTGKRTPELANNVA 893  
Qy 898 SLWFMSLFICIPATSIEMRWSGVGIDDMRNBOFWVIGVSSHFAVQGLKVIAGVD 957  
Db 894 SILFVALFSSIAITGILEMQKVGKIDDMRNBOFWVIGVSAHLFALFQGLKVIAGVD 953  
Qy 958 TSFTVTSKAGDDEEFSELYTKKTTLLIPPTLLLPFGVAVASNAINGYSGWGLF 1017  
Db 954 TNFTVTSKADDDGSESLYIFKMTSLIPPTLLIINIVGIVGVSDAISNGYDSWGLF 1013  
Qy 1018 GKLFPAPFVIVHLVPLKGLVGRNRTPTIYVMSILLASIFSLMWRIDPFALKDDEPL 1077  
Db 1014 GRLEFPALMIVHLVPLKGLGKODRVPITILVMSILLASITLLMVRVDPFVAK-GGPI 1072  
Qy 1078 LEECGGLDC 1085  
Db 1073 LEICGLDC 1080

RESULT 4  
T51579  
cellulose synthase catalytic subunit (IRX3) - Arabidopsis thaliana  
M.Alternate names: protein T1086 80  
C.Species: Arabidopsis thaliana (mouse-ear creas)  
C.Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C.Accession: T51579  
R.Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; M  
submitted to the Protein Sequence Database, August 2000  
A.Reference number: Z25394  
A.Accession: T51579  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-1026 <SAT>  
A.Cross-references: EMBL:AL391142  
A.Experimental source: cultivar Columbia; BAC clone T1086  
C.Genetics:  
A.Map position: 5  
A.Introns: 25/3; 91/1; 184/1; 215/2; 304/2; 419/3; 507/3; 578/3; 650/1; 715/3; 833/3  
A.Note: T1086\_80

Query Match 67.0%; Score 3967.5; DB 2; Length 1026;  
Best local Similarity 66.8%; Pred. No. 6.1e-236;  
Matches 736; Conservative 129; Mismatches 145; Indels 91; Gaps 16;

Qy 1 MEASAGLVAGSHNNELVIRDDPGPKPREQNGQVCOICGDDVGLAPGDPFVACNE 60  
Db 1 MEASAGLVAGSHNNELVIRDDPGPKPREQNGQVCOICGDDVGLAPGDPFVACNE 58  
Qy 61 CAFFVCRDCEYERREGTONCPCKTRRYKRLKGCQVVTGDEEDGVDDLNEFMMDGHS 120  
Db 59 CGPACRCYCYERREGTONCPCKTRRYKRLKGSFVGVGDEEDDIDDIETEFIEF-HQ 117  
Qy 121 --QSVASMLYGHMSYGRG--GDPNGA-POAFOLNPNVPLITNGQVNDIPPEQHALVP 174  
Db 121 --QSVASMLYGHMSYGRG--GDPNGA-POAFOLNPNVPLITNGQVNDIPPEQHALVP 174

```

Db      118 DKHKSABAMLYGKMSYGRGDEDEGRPP-----PVIAGHSGE----- 157
Qy      175 SFWGGG-----KRTHPLPYADPSLPVQPRSMPSKDLAAAYGGSVANKEREMNK 225
Db      158 -FPVGGGYNHGHGLKRVHPY-----PSEAGSBS-----GMRERHDDMK 197
Qy      226 QROERHQTNDGGDDGDADLPLMDARQOLSRKIPLPSSQINPYKMTIIRLVYGF 285
Db      198 L-----QHGNLGPEDD-DDPEWGLDEARQPLSRKVPPLASSKINPYRVAVIYLLAV 250
Qy      286 FPHYRVMPVNDAPALMLSVICELFWASMLDQPPKMPLEREYVRLSLRPDKGQ 345
Db      251 FRTYRLNPNVDHGLMLTSVLCIWPFAVSMILDQPPKMPLEREYVRLSLREREGE 310
Qy      346 PSQALIDFEFVSTVDLKEPPLVTNTVLSISVDPVNVKVCYVSDDAAMLTFEALSE 405
Db      311 PNMIAVADVFSVSTVDLKEPPLVTNTVLSISLAMDPEVKISCTYVSDDAAMLTFEALSE 370
Qy      406 TSEFAKKWVPCKRYNIEPRAPBWPFOOKIDYLDKRVANFVRERRAMKREYEEFKVIN 465
Db      371 TAEFAKKWVPCKKFSIEPRAPBWPFTLKVDYLDKRVHPTFVERERRAMKREYEEFKVIN 430
Qy      466 ALVAKAOKVPEBGMWODGTPMPGNVNRDHPGMIOVFLGSGGLDCEGNEPLRLVYVSE 525
Db      431 AOVAKASKVPLEGMWODGTPMPGNNTKQHPGMIOVFLGSGGFVGEHELPLVYVSE 490
Qy      526 KRPVYHHKKAGAMNALVRSAVLTNAPYLTNADCHYINNSKAIKEAMCFMMDPLLGK 585
Db      491 KRPFGHKKAGAMNALVRSAVLTNAPYLTNADCHYINNSKAVAEAMCFMMDPLDQIGK 550
Qy      586 VCVYQPPORPDGIDRDHRYANRNVVFDINMKGLDGIQGPITYVGTGCVFRQALVYDAP 645
Db      551 VCVYQPPORPDGIDRDHRYANRNVVFDINMKGLDGIQGPITYVGTGCVFRQALVYDAP 610
Qy      646 KTKKPPSRKNCMPKRCFCCGCCGGRKKOKTKTKPKTBKKKLFPKKEBNSPAYALGEID 705
Db      611 KGPKRKMTS-----CGCCPCGGRKKOKTKTKPKTBKKKLFPKKEBNSPAYALGEID 649
Qy      706 EAAPAEANEKAGIVNOOKLEKKFGQSSVFTSTLLENGGTLKASAPSLKEAIAHVISCG 765
Db      650 ----GABGDKHEMLSEMEPEKTFQSSIFVTSTLMEGGVPESSPAVLKELAHVISCG 705
Qy      766 YEDKTDWKEIGNIYGSVTEIILTFKMGCHGRSICYIIPKRVAFKGSAPLNSDRLOV 825
Db      706 YEDKTDWKEIGNIYGSVTEIILTFKMGCHGRSICYIIPKRVAFKGSAPLNSDRLOV 765
Qy      826 LRAALGSIKIFPSNHCPLWGY-GGGLKFLERFSYINSLVYPTSTIPLAYCTLPACIL 884
Db      766 LRAALGSIKIFPSNHCPLWGY-GGGLKFLERFSYINSLVYPTSTIPLAYCTLPACIL 825
Qy      885 TSGFITPELNNVSLWFMSLFICIFATSILEMMSGIVGIDMMRNQOFWYIGVSSHLP 944
Db      826 TDKFIMPPISTPSSLFISLFSIYTGILFELMSGVSIEMWRNQOFWYIGVSSHLP 885
Qy      945 VFOGLLVKVIAGVDSFTVTSKGGDESESLYTFKWTLLIPPTLLINPLFGVAVGSN 1004
Db      886 VVOGLLVKVIAGVDSFTVTSKATDDDFGLYAFKWTLLIPPTLLINPLFGVAVGSN 945
Qy      1005 AINNGVESMGPLFGKIFPAFWVYVLYLPFLKGVGNRPPTVIVYMSILIASIFSLWV 1064
Db      946 AINNGVQSMGPLFGKIFFSFWVYVLYLPFLKGMGRQNPPTVIVYMSILIASIFSLWV 1005
Qy      1065 RIDPFLAKDGPFLLESCGLDC 1085
Db      1006 RIDPFLVTKKGPDTSKGICNC 1026

```

```

RESULT 5
T05351
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana
N:Alternate names: protei n F8B4.110
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C/Accession: T05351
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Daseville, R.; De Clerck, R.; D
eues, H.W.; Mayer, K.F.X.; Schaeffer, C.
submitted to the Protein Sequence Database, February 1999
A:Reference number: 215409
A/Accession: T05351
A/molecule type: DNA
A/Residues: 1-1081 <BEV>
A/Cross-references: EMBL:AL034567
A/Experimental source: cultivar Columbia; BAC clone F8B4
C/Genetics:
A:Gene: RSW1
A:Map position: 4
A:Introns: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1;
A>Note: F8B4.110
C/Keywords: glycosyltransferase; hexosyltransferase; P-loop

Query Match      66.7%; Score 3945.5; DB 2; Length 1081;
Beet Local Similarity 66.9%; Pred. No. 3.2e-294;
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;

Qy      1 MEASAGLVASGHNRELTVIRDDGPPGKPREQNGOVCOICGDDVGLAPGDPAVACNE 60
Db      1 MEASAGLVASGHNRELTVIRHSDGKTPDKMMNGQICQICGDDVGLAETGDVFAVACNE 60
Qy      61 CAFVYRDCYERREGTQNCPOCKTRYKLGQCVYTGDEEDGVDDLNEEFMWDGDS 120
Db      61 CAFVYRDCYERREKDGQCCPOCKTRFRHGRGSRVGEDEDDVDIDENEFNY----A 116
Qy      121 QSVASMLYGMSTYRGCDPNGAPQAPOLANNVPLITNGQVVD---IPPGQHALVSPF- 176
Db      117 QGANVA-----RHQRHGEBSRSSRHSQP-IPLTGHTVSGEIRTPDQSVRTTSGP 169
Qy      177 MGGGGRTHPLPYADPSLPVQPRSMPSKDLAAAYGGSVANKEREMNKROER--HQGT 234
Db      170 LQPSRNALISSPYIDPRPVRYIVDSKDLNSIGLVNMDKKEVGGKLOKQKMLQMT 229
Qy      225 GN-----DGGDDGDADLPLMDARQOLSRKIPLPSSQINPYKMTIIRLVYGF 285
Db      230 GKYHGKGGEIEGTGSNGE--ELQWADDTRLPMKSRVVDIPSSRLTPYRVVILRLILCF 287
Qy      286 FPHYRVMPVNDAPALMLISVTCIWFAMSMILDQPPKMPLEREYVRLSLRPDKGQ 345
Db      288 FLQYRTTHPVNKAAPLMLTSVICIWFAMSMILDQPPKMPLEINREYVRLSLAIRDRGB 347
Qy      346 PSQALIDFEFVSTVDLKEPPLVTNTVLSISVDYPVDKXSCVYSDGAAMLTFEALSE 405
Db      348 PSQALIDFEFVSTVDLKEPPLVTNTVLSISVDYPVDKXSCVYSDGAAMLTFEALSE 407
Qy      406 TSEFAKKWVPCKRYNIEPRAPBWPFOOKIDYLDKRVANFVRERRAMKREYEEFKVIN 465
Db      408 TAEFAKKWVPCKKFSIEPRAPBWPFTLKVDYLDKRVHPTFVERERRAMKREYEEFKVIN 467
Qy      466 ALVAKAOKVPEBGMWODGTPMPGNVNRDHPGMIOVFLGSGGLDCEGNEPLRLVYVSE 525
Db      468 ALVAKAOKVPEBGMWODGTPMPGNNTRDHPGMIOVFLGSGGLDCEGNEPLRLVYVSE 527
Qy      526 KRPVYHHKKAGAMNALVRSAVLTNAPYLTNADCHYINNSKAIKEAMCFMMDPLLGK 585
Db      528 KRPFGHKKAGAMNALVRSAVLTNAPYLTNADCHYINNSKAIKEAMCFMMDPAIGK 587
Qy      586 VCVYQPPORPDGIDRDHRYANRNVVFDINMKGLDGIQGPITYVGTGCVFRQALVYDAP 645
Db      588 CCVYQPPORPDGIDRDHRYANRNVVFDINMKGLDGIQGPITYVGTGCVFRQALVYDAP 647
Qy      646 KTKKPPSRKNCMPKRCFCCGCCGGRKKOKTKTKPKTBKKKLFPKKEBNSPAYALGEID 705
Db      648 LTER-----DLEPNIIVKSCGSRKKGKSSKKNYERKRI--NRSQSNAPLPMEDID 699
Qy      706 EAAPAEANEKAGIVNOOKLEKKFGQSSVFTSTLLENGGTLKASAPSLKEAIAHVISCG 765
Db      700 EBFEGYDDERSTLMSQSRVEKRFQSPVFIATPFGGIDPPTNPATLLKEAIAHVISCG 759

```



Qy	766	VEDKTDKGEJGIMIGSVLTEDILTGPHOMCHQMBSIYCCPKGVAFKSGAPLNLSDRLHOV	825
Db	760	YEDKTWGEJGIMIGSVLTEDILTGPHQARMGIMISYCNPPRPAFGSAPINLSDRLNOV	819
Qy	826	LRMALGSIEIFESNHCPLMWYGGGLKFLERESYINSIYVPTSIPFLAYCTLPALCILT	885
Db	820	LRMALGSIEILISRCPIWYGHGRLLERLAVINTIYPTISIPILAYCILPACILTT	879
Qy	886	GKFIIPBLNNVASLWFMSLFCIFATSIILEMBSGVGIDDMWRNEQFWYIGGVSSHLPAY	945
Db	880	DRFIIPBLISNVASIMFILFFISIAVTGILELMSGVSIEDMWRNEQFWYIGGVSAHLFPAY	939
Qy	946	FOGLILKVIAGVDSTSTVTSKGD-DEESELVTFKWTLLIIPPTLLILNFIIGVAVGSN	1004
Db	940	FOGLILKVIAGIDTNTFTVTSKATDEBDGFALVITFKMTALLIPTVILVNLIGIVAGVSY	999
Qy	1005	AINNGYBESNGPLFGCLLFPAFWYIVHLYPFLKGLVGRQNTPTIIVYMSILLASISFELWV	1064
Db	1000	AVNSGYQSGWGPLFGLLFFALWVIAHLYPFLKGLGRQNTPTIIVYMSVILASISFELWV	1059
Qy	1065	RIDPFL 1070	
Db	1060	RINPFV 1065	

## RESULT 6

Cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 31-Dec-2000  
C:Accession: J152054  
R:Arnold, T.; Peng, L.; Betzner, A.S.; Burr, J.; Wittke, W.; Herth, W.; Camilleri, C.; F  
Science 279, 717-720, 1998  
A:Title: Molecular analysis of cellulose biosynthesis in *Arabidopsis*.  
A:Reference number: Z13745; MUID:98111412; PMID:9445479  
A:Accession: T52054  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1065 <ARI>  
A:Cross-References: EMBL:AF027174; PIDN:AA039336.1  
C:Genetics:  
A>Note: Atn-B  
C:Function:  
A:Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; involved in ac  
Keywords: glycosyltransferase; hexosyltransferase

Query Match	65.5%	Score 3875.5	DB 2	Length 1065
Best Local Similarity	66.7%	Pred. No. 7.4e-289		
Matches	728	Conservative 145	Mismatches 161	Indels 57
Gaps				18
QY	23	DGDPRKPRBQNGVCQICGDDVGLABGDFVACNCAFPVCRDCEYERRECTQNCP	82	
Db	4	EGETTAGKMKMKIIVPQTQCISDDNVGKTVGDRFVACDICSFPVCPCPYEYERKDNQSCP	63	
QY	83	QCKRKYRKLKCCQARPTGDEEDGV-DLDLNERFMNGHDSQSAVESMLVGHMSYGCGGPN	141	
Db	64	QCKRKYRKLKCSPRIPGDKDEGLADGCTVEKNYP-QEKESISMLGMHLIRGL-GBEM	120	
QY	142	GAPQAFQ-LNPVNPVLITNGQVWDIIPRCHALVPSFMG-----GGGRIRHPLPY-ABPS	193	
Db	121	GEPYDKKEVSHNHLPRLTSRQ---DTSGEFSALSPERLVSSTINGKRK-LPYSQVN	174	
QY	194	LPVQPRSMDSKDLAAVGYGSVAMERBMENKKORER-----MHGTGNDGGDDGDD---	245	
Db	175	QSPNRRIYD-----VGLGNVAMERADVGMMKEKNTQPGVSTQPAASRGCVDDIDASTD	228	
QY	246	--ADLPLM-DEARQOLSRKIPLPSSQINPYMKIIIRLVVLGFPPHYRVMHPVNDAFALM	302	
Db	229	ILADEALLNDEARQGLSRKVSIPSSRINPYRNVIMLRVLVILCLFLHYVITNPVPAPFLM	288	
QY	303	LISVTCETWFSAMSWLDPFKMFLERETYLDRSLRFDKSGQSPQLAIDFFVSTVDPL	362	

```

Db      289 VSVIYCEIMFLSMIILDOQPKMFPVNRETYLDRLALAYDRBEGBSQOLAANDIFVSTYDPL 348
Qy      363 KEPLVLTNTVYLSILSDVDPYDKVSCVSDGAAMLTFEALSETSEFAKTMVPCRYNI 422
Db      349 KEPLVLTANTVLSITLAVDPYDKVSCVFDGAAMLSFEISLAETSEFAKTMVPCCKXYSI 408
Qy      423 EPRAPENYFOOKIDLYLKDKAANAVRRRRAMKREYEEFKRINALYAKAQVPEEGTMQ 482
Db      409 EPRAPENYFAKIDYLDKQTSFVKDRRRAMKEYEFEFKRINALVSKALKEPBEQVMQ 468
Qy      483 DGTMPGNGNVRDHGMIQVFLQSGGJDCGAGNELPRLVYVSRERPCYNNHKKAGANAL 542
Db      469 DGTMPGNGNNGDHGMIQVFLQSGGJDCGAGNELPRLVYVSRERPCGFORHKKAGANAL 528
Qy      543 VRSAVLTNAPYLLNLDCDHYINNSKAIKEAMCPMDPLGKKVCYQVQPORFGDIRHD 602
Db      529 VRSAVLTNBPFIINTLCDHYINNSKALREAMCTLMDPNLQKOVCIYQVQPORFGDIRKD 588
Qy      603 RYANRNVVPDINNKGLDGIQGPITYGTCVFRQALXYDAP---KTKPP--SRTCNC 657
Db      589 RYANRNVVPDINNRGLDGIQGPITYGTCVFNRTALXYEPRPIKVMHKKPSSLKSLCG- 647
Qy      658 WPKMCFCCCCGNGRKQKTKTKTEKKLLPFKKENQSPAYALGEIDEAPGA--ENEK 715
Db      648 -----GSRKNSKAKKESDDKK--SGRHTDSTVPVFNLDIEBCVGAQFDEK 694
Qy      716 AGIYNQCKLEKKFQSSVFTSTILNENGFTLKASPASSLKEAIVHSICGYEDKTMGKS 775
Db      695 ALLMSQMSLERFQQSAVFVASTLMEGQVPPSATPENLKEAIVHISICGYEDKDWGME 754
Qy      776 IGMTYGSATBEDILGFPMHCHGMRSYICIPKRVAFKGSAPLNSDRLHOUTRLMAGSIEI 835
Db      755 IGMTYGSATBEDILGFPMHARGMSYICMPLPFPKASAPINISDRNLQVLRALAGSVEI 814
Qy      836 PFSNHCPLMYGGGGLKFLERFSYINSIVPMTSIPILAACTLPALCLTLGKFTPTPLN 895
Db      815 LFSNHCPIWICYNKRLLFLERFAVNTTIYIPIISIPILMTCITLAAVCLFTNQPITPOISN 874
Qy      896 VASLMEFMSLFCIPATSILEMRSVGVIDDMWRNEQCFWVIGVSSHLEFAVFOGLIKYAG 955
Db      875 IASLMPISLFLSATGILEMRSGVGIDEMWRNEQCFWVIGVSAHLFAVFOGLIKYAG 934
Qy      956 VDISFPTYSKGD--DEEPSLYTFKMTLLIIPPTLLILNFIGVAVGSAINNGYSWG 1014
Db      935 IDTFPTYSKASDDEGDFAEIYLFKMTLLIIPPTLLIIVNLGVAVGSVAINSQYWG 994
Qy      1015 PLFEGLEFAFVVIYHLVPLKGLVGRONRPTIYIWSIILASIFSLKATRIDPLAKD 1074
Db      995 PLFEGLEFAFVVIYHLVPLKGLMGRONRPTIYIWSVLLASIFSLIMRIDPFTSRV 1054
Qy      1075 GPILLEGGDLC 1085
Db      1055 GPDLIEGGINC 1065

```

## RESULT 7

C:probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:date: 02-Feb-2001 #sequence\_rev:0102-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: F84649  
 R:Rilln, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujiki,  
 M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.B.; Umayam, L.; Ta-  
 euse, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; V  
 Nature 402, 761-768, 1999  
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: F84649  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1065 <STO>  
 A:Cross-references: GB:AE002093; NID:g4432865; PIDN:AAD20713.1; GSPDB:GN00139  
 C:Genetics:

A:Gene: At2g25540  
A:Map position: 2

Query Match 63.9%; Score 3782; DB 2; Length 1065;  
Best Local Similarity 63.9%; Pred. No. 1,1e-281;  
Matches 702; Conservative 150; Mismatches 168; Indels 78; Gaps 15;

```

QY 7 LVAGSHNNELVIRRDGPPKPPREQNGVQICGDDVGLAPGGDPVVAQNECAPYVC 66
DB 1 NVAGSVRRYFVRNDDSDGKPLKDLNGOICQICGDDVGLKTKGNVFAQNECGFPILC 60
QY 67 RDCYERREGEONCPCKCTRYKRLKGCORVTGDEEDVDLQDNBFNMDGHSQVAES 126
DB 61 QCTEYERKDSQCCQCKARFRNRANGSRVEYDEKEDVDNIENEFD----- 108
QY 127 MLYGHMSYGRGSDPNGAP---QAFQINPN-----VPLLTNGQVVD-DIP-DEQHALVPS 175
DB 109 -----YTGNNKARLPHRABEFSSSHNEESLPVSLTTHGHPVSGELPTDRNATLS- 160
QY 176 FMGGGGRKTHPLPYADPSLP-----VQPSMDPSKDLAAYGGSVANKEREMWKKORQ 228
DB 161 -----PCIDPQLPGIYQLLLFVRILDPKIDNSYGLVNVVDMKKRIQGWKLMQ 208
QY 229 ER--WQGTGN-----DGGDDGDDADLPLMDARQQLSRKILPSSQINPYRMII 277
DB 209 DKQMTHTMGTGYHEGKGEGEGEGSND--ELQVNDARLPMKSRVHFFSARMTPTVIV 266
QY 278 IRLVVGFFPFRVVMVNDAPALMLISYCEIWFMSWILIQPFKPFIEETYLDRLS 337
DB 267 LRLIILIGVLRHTTHPVDAVALMLTSVICEIWFMSWILIQPFKPFIEETYLDRLS 326
QY 338 IRLPDGQPSQLAPIDFVSTVDPLKEPLVTTNTVLSLSVDYVDKVSQVSDGAM 397
DB 327 LHYDRGEGSQLAPVDVFTVTDPMKEPLVTANTVLSLAVDYPDKAACVYSDGSM 386
QY 398 LRFPAISESEFAKKNVPRCKRYNIEPRAPEWFOQIDYLDKQVAAVRRBRAMKRY 457
DB 387 LRFPAISESEFAKKNVPRCKRYNIEPRAPEWFOQIDYLDKQVAAVRRBRAMKRY 446
QY 458 EEFKVRINALVAKAQVPEEGWMDQTPMPCNNVNDHGMIOVFLGSGSGLDCEGNEPL 517
DB 447 EEFKVRINALVAKAQVPEEGWMDQTPMPCNNVNDHGMIOVFLGSGSGLDCEGNEPL 506
QY 518 RLIVYSREKRCGYNHKKAGAMNALVRSVAVLTNADYLLNDCDHYTNNSKAIKEAMCM 577
DB 507 RLIVYSREKRCGYNHKKAGAMNALVRSVAVLTNADYLLNDCDHYTNNSKAIKEAMCM 566
QY 578 NPLBLGKVCYVQPRPFGIDRHDRVANRVFPIINNGKLDGICQIPYVGGCVPRRO 637
DB 567 NPLBLGKVCYVQPRPFGIDRHDRVANRVFPIINNGKLDGICQIPYVGGCVPRRO 626
QY 638 ALVGYDAPRTKK--PSPRTCNCPKMKCFCCCFGNKOKKTTK-PKTEKKKLLFFKKEBN 694
DB 627 ALVGYDAPRTLEEN-----IYKSCFGSKKKKSKRIPIYEDNRSL--KRSLS 675
QY 695 GSPAYALGRIDEAPAEENEKAGIVNOQLEKKEGSSVPVSTLLENGCYLKSASPAL 754
DB 676 NPLLFNMBIDIDEDVEGEDEMSLLVSKRLERKFGQSPVFIATFMEQGLPSTTLP 735
QY 755 LKEATHIVISCGYEDKTDKMGKEIGWITGSVTEEDILTFPKHCHGMRITICLPKXVAKGSA 814
DB 736 LKEATHIVISCGYEDKTDKMGKEIGWITGSVTEEDILTFPKHCHGMRITICLPKXVAKGSA 795
QY 815 PLNLSDRLQVLRMALGSIIEIFPSNHCPLVYGGGLKFLERPSYNSIVYPTISPLLA 874
DB 796 PNLSDRLQVLRMALGSIIEILSRHCPIWYNGSLKLLERLAIYINTIYPTISPLLA 855
QY 875 YCTLPALICLTGKFTPELNNVSLMFMISFICIFATSIIEKMSGVGIDDMWRNQFV 934
DB 856 YCTLPALICLTGKFTPELNNVSLMFMISFICIFATSIIEKMSGVGIDDMWRNQFV 915
QY 935 IGGVSHLFAVFPQGLKVLAVGVTSFTVNSKGD--DEPSELTFFKMTLLIPPTLLLL 993
DB 915 IGGVSHLFAVFPQGLKVLAVGVTSFTVNSKGD--DEPSELTFFKMTLLIPPTLLLL 933

```

```

DB 916 IGTSAHLFAVFPQGLKVLAVGVTSFTVNSKGDDEGDPALVYFKMTSLIIPPTLLLV 975
QY 994 NFIGVAGVSNAINNGESWGPLFGKLFPAFWIVLHLYPLKGLVGRQNTPTTVIWSI 1053
DB 976 NLVGVAGVSNAINNGESWGPLFGKLFPAFWIVLHLYPLKGLVGRQNTPTTVIWSA 1035
QY 1054 LLAISFSLWVRIDPFLA 1071
DB 1036 LLAISFSLWVRINPVS 1053

```

# RESULT 8

T10797  
cellulose synthase (EC 2.4.1.-) catalytic chain cel1 - upland cotton  
C:Species: Gossypium hirsutum (upland cotton)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10797  
R:Pear, J.R.; Kawagoe, Y.; Schreckengost, W.B.; Delmer, D.P.; Stalker, D.M.  
A:Title: Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase  
A:Reference number: Z17152; MUID:97057296; PMID:8901635  
A:Accession: T10797  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-974 <PEA>  
A:Cross-References: EMBL:U58283; NID:G1706955; PID:AA837766.1; PID:G1706956  
A:Experimental source: strain Acala SU-2; fiber  
A:Genetics:  
A:Function: cel1  
A:Description: involved in the synthesis of cellulose  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 58.6%; Score 3469.5; DB 2; Length 974;  
Best Local Similarity 59.8%; Pred. No. 9.6e-258;  
Matches 638; Conservative 152; Mismatches 168; Indels 109; Gaps 13;

```

QY 33 EONGVQICGDDVGLAPGGDPVVAQNECAPYVCRCYERREGEONCPCKTRYKR-- 90
DB 3 EGVVVCCHTCEHVGALNNGEPFVACHCNPICKSCPEYDLKEGRKACLCGSGYDENL 62
QY 91 LKCGORVYDEEDVDLQDNBFNMDGHSQVAESMLYGHMSYGRGSDPNGAPQAFQIN 150
DB 63 LDDVEKATGD-----QSTAAHLNKSQD----- 85
QY 151 NPLVPLTNGQVVDIPRQHALVPSFMGGGGRKTHPLPYADPSLPVQPSMDPSKDLAY 210
DB 86 -----VGIHARHISVSITLSE-----AB 105
QY 211 GYGSVAMKERNMKQROERHMQTGNDGGDDADLP---LMD---EARGQLSKTI 262
DB 106 DNGNSIMKQRYESWKEKKKKKKPAT--KYEREAELIPRQOMEDKPAAPASQPLSTII 162
QY 263 PLSSQINPYMIIIRIVLGFPPFRVVMVNDAPALMLISVCEIWFMSWILIQPFKPF 322
DB 163 PLSSQINPYMIIIRIVLGFPPFRVVMVNDAPALMLISVCEIWFMSWILIQPFKPF 222
QY 323 KMFPIERTYLDRLSLRFDEKQPSQLAPIDFVSTVPLKEPLVTTNTVLSLSVDP 382
DB 223 KMFPIERTYLDRLSLRFDEKQPSQLAPIDFVSTVPLKEPLVTTNTVLSLSVDP 282
QY 383 VDKVSCYSDGAAALTEALSETSEFAKKNVFPCKRYNIEPRAPEWFOQIDYLDKXV 442
DB 283 VDKVSCYSDGAAALTEALSETSEFAKKNVFPCKRYNIEPRAPEWFOQIDYLDKXV 342
QY 443 AANFRRERAMKRYEERFVKRINLVAQKQVPEEGWMDQTPMPCNNVNDHGMIOV 502
DB 343 AANFRRERAMKRYEERFVKRINLVAQKQVPEEGWMDQTPMPCNNVNDHGMIOV 402
QY 503 LGQSGGLDEGNEPLRLVYVSREKRCGYNNHKKAGAMNALVRSVAVLTNADYLLNDCDH 562
DB 403 LGQSGGLDEGNEPLRLVYVSREKRCGYNNHKKAGAMNALVRSVAVLTNADYLLNDCDH 462

```

[illegible][illegible]

A;Molecule type: mRNA  
 A;Residues: 1-685 <PEA>  
 A;Cross-references: EMBL:U58284; NID:g1706957; PIDN:AAB37767.1; PID:g1706958  
 A;Experimental source: strain Acala S9-2; fiber  
 C;Genetics:  
 A;Gene: celA2  
 C;Function:  
 A;Description: involved in synthesis of cellulose  
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.1%; Score 2907; DB 2; Length 685;  
 Best Local Similarity 75.0%; Pred. No. 8.9e-215;  
 Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

```

QY 411 KKWPFCGRYNIIEPRADWFWFOQKIDYKDKVANFPRRRAMREYEFKRVINALVAK 470
   :::::::::::::::::::::
Db 1 RRVVFCGRKINVERAPREFYNEKIDYKDKVHSPFKERRAMREYEFKRVINALVAK 60
   :::::::::::::::::::::
QY 471 AQRVBEQWTDGTPMPGNVVDHFGMIQVFLGSGGLDCEGNEPLRVLYVSREKRGY 530
   :::::::::::::::::::::
Db 61 AQRKPEEGWWDGTPMPGNVVDHFGMIQVFLGSGALDVGKELPRLVYVSREKRGY 120
   :::::::::::::::::::::
QY 531 NHHKKAAMALVAVSVLTNAPRLVLDCHYINNSKAIKEMCFMMDPLGKVCYVQ 590
   :::::::::::::::::::::
Db 121 QHHKKAAGANALVAVSVLTNAPRLVLDCHYINNSKAMREACFLMDPQFGKCYVQ 180
   :::::::::::::::::::::
QY 591 PPRFGIDRDRYANNNVFPDINMGDLGDIQPIYVGTGCVFRQALGYDAPKTKR 650
   :::::::::::::::::::::
Db 181 PPRFGIDRDRYANNNVFPDINMGDLGDLQSPVYVGTGCVFRQALGYDAPVSEKR 240
   :::::::::::::::::::::
QY 651 PSRTCNCPKWCFCGCCGFGNRKQKTKPKTEKK-----KLLFPKK-----EENQSP 697
   :::::::::::::::::::::
Db 241 PKMTCDGMPWC-CCCCGGSRKK---SKKGGKKGLGGLYGGKKMGKNNYKKGSA 296
   :::::::::::::::::::::
QY 698 AYALGEIDAPAKEN-EKAGIVNQQLEKKFGGSSPVVSTLLENGTLKASPA 756
   :::::::::::::::::::::
Db 297 VFDLEIEEGLEGEYELKSTLMSQKNFEKFGSPVFIASITLMENGLPEGTNSTLIK 356
   :::::::::::::::::::::
QY 757 EAHIVISGVEDTKDMKEIGMIVGSTEDILTFGKHCGRMSIYICIPKRVAFKGSAPL 816
   :::::::::::::::::::::
Db 357 EAHIVISGVEKTEKEMKEIGMIVGSTEDILTFGKHCGRMSIYICIPKRVAFKGSAPL 416
   :::::::::::::::::::::
QY 817 NLSDRLHQVLMALGSIIEFPSNCHPLMYGGGGLKLEPERSYINSIVYPTWSTPLLA 876
   :::::::::::::::::::::
Db 417 NLSRLHQVLMALGSIIEFPSNCHPLMYGGGGLKLEPERSYINSIVYPTWSTPLLA 476
   :::::::::::::::::::::
QY 877 TLPAICLTGKFTPELANNVSLMPSLFCIFATSTLEBRMSGVGIDDMRNEQFV 936
   :::::::::::::::::::::
Db 477 TTPAVCLLTGKFTPTISNLTSMVFLALFSLIATGVLBRMSGVSIQDMWRNEQFV 536
   :::::::::::::::::::::
QY 937 GVSRLFAVFGGLKLVAGVDTSTFTVYSGKGDDEFSFLYFKWTLLIPTTLINFI 996
   :::::::::::::::::::::
Db 537 GVSRLFAVFGGLKLVAGVDTSTFTVYSGKADDTFEGELVFKWTLLIPTTLINFI 596
   :::::::::::::::::::::
QY 997 GVAVGSNAINGEYSGPLFGKLFPAFWIVLHLYPPLKGLVGRQNRPTIVVMSILA 1056
   :::::::::::::::::::::
Db 597 GVAVGSDAINGEYSGPLFGKLFPAFWIVLHLYPPLKGLVGRQNRPTIVVMSILA 656
   :::::::::::::::::::::
QY 1057 SIFSLVWRIDPLAKODGPLBECGLDC 1085
   :::::::::::::::::::::
Db 657 SIFSLVWRIDPLPKOTGPVLKOCGVGC 685
   :::::::::::::::::::::

```

## RESULT 11

hypochemical protein F22D16.26 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C;Accession: D86157  
 R;Theologis: A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon  
 ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: D86157  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1181 <STO>  
 A;Cross-references: GB:A8005172; NID:g6056428; PIDN:AAF02892.1; GSPDB:GN00141  
 C;Genetics:  
 A;Function:  
 A;Keywords:

A;Map position: 1

Query Match 39.4%; Score 2310.5; DB 2; Length 1181;  
 Best Local Similarity 43.0%; Pred. No. 3.5e-170;  
 Matches 489; Conservative 166; Mismatches 274; Indels 207; Gaps 26;

```

QY 31 PREQNGVQICGDDVGLAPGDPRVACNECAFVYCRDCEYERREGTQNCPOCKTRYR 90
   :::::::::::::::::::::
Db 154 PEKSGQICMLKGCDEKVVHG-----RC-EGCFRICDCEYFCITISGGGNCGCKEPRVD 207
   :::::::::::::::::::::
QY 91 LKGCQRTVGDDEBDGYDLDNEFNMGDHDSQVABSMLYGHMSYGRGDPNGAPQAPQLN 150
   :::::::::::::::::::::
Db 208 INDDPETEEDEDEDEAKPL-----PQMGSEKLDKRLS----- 239
   :::::::::::::::::::::
QY 151 PNVPLLTNQWVDIPRQHALVPSFMGGGGRINPLPYADSLPVQPSMPSKDL----- 207
   :::::::::::::::::::::
Db 240 -----VKSFKAON-----QAGDPHTRWLFET 262
   :::::::::::::::::::::
QY 208 -AAYGYSVAMKERMENKQROBRHQGT---NDGCGDDADALPLMEARQQLSRKIP 263
   :::::::::::::::::::::
Db 263 KGTGYGNAVMP-----KQYIGSGGGGNGYETPRPFGSRKSRPLTRKYS 308
   :::::::::::::::::::::
QY 264 LPSSQINPYRMIIIRLVVLFPPHYRVNHPVNDAPALMISVTCIFAMSWILDQPK 323
   :::::::::::::::::::::
Db 309 VSAIISPYRLIARLALVGLFTWRVHPVREAMMLMGMSSTCETALMSLWLDQLPK 368
   :::::::::::::::::::::
QY 324 WPIRETEYLDRLSLRFDK-----EGQSQALPIDFSTVDPKKEPLVNTNTVLSL 377
   :::::::::::::::::::::
Db 369 LCPVNLTLGVKKEFESPNLRNPKR-SDLPGLDVFSTADPEKEPLVANTNTLSL 427
   :::::::::::::::::::::
QY 378 SVDPYVDKVCYVSDGAMLTFEALSETSEFAKWPVPCSKYNIERAPAWFOQKIDY 437
   :::::::::::::::::::::
Db 428 AVDPYVKAACYISDGGALITFEALAQVASTVWPCRGKNIERAPAWFOQKIDY 487
   :::::::::::::::::::::
QY 438 LKQKVANVRRERAMKREYEFKRVINAL-----VAKQK----- 473
   :::::::::::::::::::::
Db 488 LKQKVANVRRERAMKREYEFKRVINAL-----VAKQK----- 547
   :::::::::::::::::::::
QY 474 -----VPEEGTMDGTPMPG-----NNVRDHPMIOVFLGSGGLDCEGNE- 515
   :::::::::::::::::::::
Db 548 NPOETIVIPKATV-KSDGSHWPETMSSGETDSRGHAGIQMLAPRPAAPVYGAEADA 606
   :::::::::::::::::::::
QY 516 -----LPRLVYVSREKRPFGYNNHKKAGANALVAVSVLTNAPRYLLINDCDHY 563
   :::::::::::::::::::::
Db 607 ENLIDTVDIRLPMIVYVSREKRPYDINKKAGANALVAVSVLTNAPRYLLINDCDHY 666
   :::::::::::::::::::::
QY 564 INNSKAIKEMCFMMDPLGKVCYVQPPQRFQDIDRDRYANNNVFPDINMGDLGDIQ 623
   :::::::::::::::::::::
Db 667 IYNSMALREGMCFMD-RGQDRICYVQFPQREGIDPNRYANNNVFPDINMGDLGDIQ 725
   :::::::::::::::::::::
QY 624 GPIYGTGCVFRQALGYDAPKTKRPSRTCNCPKWCFCGCCGFGNRKQKTKPKTKR 683
   :::::::::::::::::::::
Db 726 GPMIVTGTGCFRRTALYGSPPRATVHG-----W-----LGRRTVKSLSR---R 767
   :::::::::::::::::::::
QY 684 KLLFPKKEENOSPAYALGEIDAPAGANERAGIVNQQLEKKFGGSSPVVSTLLENG 742
   :::::::::::::::::::::
Db 768 PKAMMKKDBEVSIPIN--GEYNE-----BENDGDLESLLPRPFGNSNPFVAVAY 820
   :::::::::::::::::::::
QY 743 GGTL-----KSASPASL-----KEAIVISGVEDTKDMKEIGMIVGSTEDILTF 782
   :::::::::::::::::::::

```

Db 821 OGRLIDLOGKGNKSHRPSAGSLAVPREPLAATVAEAISVISCYEDKTEWKKVWYGS 880  
 QY 783 VVEDILTGFRMCHGHRSTIYCIPIKRVAFSGAPLINSDBLHOVLPRALGSIETFFENHCP 842  
 Db 881 VVEDVVTGIRMRNRGRSITCYVKRDAFGTAPINLDRHQVLRAWATSVEIFFSRNNA 940  
 QY 843 LMYGSGGKFLERPSYINSIVPMTSIPLAYCTLPALICLNGKFTPELNNVANSALPM 902  
 Db 941 IF--ATRRMKFLORVAFVFWGMVPTSLPIYVICLPAISLPSGQITVGSLODTFLIYL 998  
 QY 903 SLFICIFATSIEMRNSGVGIDDMWRNEQFWIVGVSSHFAVQGLLVAVIAGVDTSTFV 962  
 Db 999 SITLTLCLMLSTLEIKMSGITLHEMRNEQFWIVGTSAPPAVLQGLLVAVIAGVDTSTFV 1058  
 QY 963 TSKG---DDEFSLLYFFKWTLLIPIPTLLNLFVGVAVNSAINNGVSKQPLFG 1018  
 Db 1059 TSSSAPBEDGDDEPALYVKKMSFLWVPLTIMVMVIAVGLARTLVSPPQMSKLVG 1118  
 QY 1019 KLFFAFVAVIHLVPLFLKLVGRONTPTIYVMSILASIFSLMWIRIPLAKOD 1074  
 Db 1119 GVFFSFVWLCILYFPAKGLMGRGRVPTIYVMSGLSLITVSLMWYINPPSGKD 1174

## RESULT 12

T51546  
 cellulose synthase catalytic subunit-like protein - Arabidopsis thaliana

N.Alternate names: protein F2K13\_60

C.Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 18-Aug-2000 #sequence\_rev18-Aug-2000 #text\_change 18-Aug-2000

C.Accession: T51546

R.Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

A.Reference number: Z25394

A.Accession: T51546

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-1145 <SAT>

A.Cross-references: EMBL:AL391141

A.Experimental source: cultivar Columbia; BAC clone F2K13

C.Genetics:

A.Map position: 5

A.Introns: 297/2; 566/3

A.Note: F2K13\_60

Query Match 39.1%; Score 2311; DB 2; Length 1145;

Best Local Similarity 42.6%; Pred. No. 1e-168;

Matches 493; Conservative 170; Mismatches 271; Indels 222; Gaps 33;

QY 11 SHNRNLVYIRRDGDPKPRPRQNGOVQICGDVYGLAPG--GDFVACNECAFVCRD 68  
 Db 100 SNTRAHLMKRVLETENHPQOMAGSKSSCAIPGCDAKWMSDERGQDLPC-ECDFKICRD 158  
 QY 69 CYEYERREGTQNCPOCKTRYKRLKGCQVYTGDEEEDGVDDLDNEFWMDGHSQVASEML 128  
 Db 159 CFIIDAVKTTGGGICPGCKEKYKNT---THLTDVDENG----- 191  
 QY 129 YGMSYSGRGDPPGAPQAFQALPNVPLLTNGQVWDIPEQHALVPSFMGCGG---KRI 184  
 Db 192 -----QQRMLP-----GGGSKMERL 209  
 QY 185 HPLPYADPSLPVQPRSMDS-----KDLAAYGYSVANKEMENMKQQRKHQGTGNG 238  
 Db 210 SWKSTNKSALMSQGDGPDHNRMLTETGTGYGNAFWT-----KDGDFG 255  
 QY 239 GGDGDD-DAD-----LPLMDEARQOLSKRIPLSSQIPYNNIIIRLVVLGFFPYRV 291  
 Db 256 SKDGDGDDGDMGMEKQDLMRSRWRPLTKLKIPIAVISPIRLITRIVLALPLTWAV 315  
 QY 292 MHPVNDALMLISVLCIWMFAMSWILDFPKMFPIERTYLDRLSLRPD--KEGQP--- 346  
 Db 316 KIQNPDAVVMWGMGVSVCLELWFAISWLDQLPKLCPIRATNDLQVLEKEKETPTASNPQK 375  
 QY 347 SQLAPIDPFVSVYDPLKEPPLVTNTVLSLSVDYFVDKVSCTVSDGAAMLTFEALSET 406

Db 376 SDLPGFDFVFSVADDEKEPPLVTANTLISILAAEPVEKLSCTVSDGALLTFEAMEA 435  
 QY 407 SEFAKKWVPFCRRYNIERPAPRWYQOKIDYKDVAAANFVRERPMKREIEFFVRINA 466  
 Db 436 ASFANIWPFCRKAIEPRNPDSYSLKRDYKKNKXSDPFVDRRRVRRFEDEFKRVNS 495  
 QY 467 L-----VAKA-----KYDEGWTMQDTPMGNV--- 493  
 Db 496 LPDSIRRSADVAHAEETKAMQKQNRNDEBMEPVKIPKATW-NADGTHMGTVLTSAS 554  
 QY 494 -----DHGMQLQVFL-----GQSGG-LDCEGNE--LPLVYVSREKRGYNHKK 536  
 Db 555 DHAKGDHAGIIQVMLKPSDEPLHGVSGFLDITVDRLPLVYVSREKRGYDHAKKA 614  
 QY 537 GAMNALVRSVAVLTNAPYLINDCHYINNSKAIKAMCFMMDPLLGKVKCVYQPPORD 596  
 Db 615 GAMNALVRSVAVLTNAPYLINDCHYINNSKAIKAMCFMMDPLLGKVKCVYQPPORD 673  
 QY 597 GIDRDRYANRNVVFPDINMGDLOGIPYVGTGCVRRQALYGYDAKTKK-PPSKTC 655  
 Db 674 GIDPSDRYANRNVVFPDINMGDLOGIPYVGTGCVRRQALYGYDAKTKK-PPSKTC 730  
 QY 656 NCMPRKFCGCCFCGGRKQKTKTPKTEKKLLFFKGENOSPAYALGEIDSAAPGAENK 715  
 Db 731 -CM-----SCF-----PSKAKNI---PEENR-ALNMSDYDBE----- 760  
 QY 716 AGIVNQKLEKKFGGSSVFTVS-----TLLENG---GTL---KSASPSLL 755  
 Db 761 ---NMLSLVPRKFGNSTFLDISIPVAFQGRPLADHPAVKNGRPPGALTIPELLDAGTV 817  
 QY 756 KEAHIVICGTYDKTDNMGKEIGMIGSTEDILITFGKMGCHGMSIYCIPIKRVAFKSGAP 815  
 Db 818 AEAIAVICWEDKTEKMSRIGMIGSTEDILITFGKMGCHGMSIYCIPIKRVAFKSGAP 877  
 QY 816 INLSDRHLQVLRWALGSEIFPSNHCPLMYGSGGLKFLERPSYINSIVPMTSIPLAY 875  
 Db 878 INLTRLQVLRWALGSEIFPSNHCPLMYGSGGLKFLERPSYINSIVPMTSIPLAY 935  
 QY 876 CTLPAICLLTGKFTPELNNVASMFWMSLFCIFATSIEMRNSGVGIDDMWRNEQFWIV 935  
 Db 936 CFIPLALSLFSGQFIQTLNVTFVLYLLIISTLCALLLEIKMSGISLEBMRNRQPMHI 995  
 QY 936 GGVSSHLLAVRFGGLKLVAGVDTSTVTSK-GGD--DEFSLLYFFKWTLLIPIPTLL 992  
 Db 996 GGTSHLLAVRFGGLKLVAGVDTSTVTSK-GGD--DEFSLLYFFKWTLLIPIPTLL 1055  
 QY 993 LNFIGVAVGVSAINNGVSGPLFGKLPFAFWVVLVYPLKGLVGRONTPTIYVMS 1052  
 Db 1056 VNLIAVAGFSRTIYSVVPQMSKLVGVFSFWLALHLYPPAKGLMGRGRTPTIYVMS 1115  
 QY 1053 ILLASIFSLMWIRIDP 1068  
 Db 1116 GLVATLISLWVAINP 1131

## RESULT 13

T05646

hypothetical protein F20D10\_310 - Arabidopsis thaliana

C.Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 23-Apr-1999 #sequence\_rev18-Aug-1999 #text\_change 23-Jul-1999

C.Accession: T05646

R.Bevan, M.; Medler, H.; Kutner, M.; Wambutt, R.; Bancroft, I.; Mayer, K

submitted to the Protein Sequence Database, February 1999

A.Reference number: Z15420

A.Accession: T05646

A.Molecule type: DNA

A.Residues: 1-1111 <BEV>

A.Cross-references: EMBL:AL035538

A.Experimental source: cultivar Columbia; BAC clone F20D10

C.Genetics:

A.Map position: 4

A.Introns: 139/2; 675/3

A&gt;Note: F20D10.310

Query Match 38.1%; Score 2256; DB 2; Length 1111;  
 Best Local Similarity 41.7%; Pred. No. 1,7e-164;  
 Matches 472; Conservative 163; Mismatches 260; Indels 236; Gaps 27;

QY 36 GGVQCI--CGDDVGLAGDPPVACNECAFVPCRCDCYERREBGTONGCQCTRYKRLG 93  
 DB 112 GSSCAMPACDGVNMKDERGKDVPC-ECRFKICRDFMDAKE-TGLCFGCCKEYKX---- 165  
 QY 94 CCRVUGDEBEDVDLDNENFMDGHSQSVASBSMLYGHMSYGRGDPNGAPQAFOL--NP 151  
 DB 166 ----IGDDDDTDP-----YSSGALPLPAPGKQGRANN 195  
 QY 152 NVPLLTNGMVDDIPPEQHAYPSFMGGGGRTHPLPYADPSLPVQPRMSDPSKDL---- 207  
 DB 196 NMSMKRNC-----NGE-----FDHNRMLFETQ 218  
 QY 208 AAYGYSVAMKERMENKOROEKMHOTGDDGGDDDD--ADLPMBARQOLSRKIP 263  
 DB 219 GTYGNAYWP-----ODEMY-----GDDMDGMRGGMVETADKPRKPLSRIP 262  
 QY 264 LBSOINPMTIIRLVVLFPEFHYRVNVPVNDALMLISVIGETWFMGMIIDQPK 323  
 DB 263 IPALITSPYRLIVFVVLCPFLTRINPNEDALMLMWSITICLWMGFSTILDQIPK 322  
 QY 324 WFERETYLDRLSLRFD--KEGQF--SQLAFIDFVSVDPLKEPPLVTTNTVLSLS 378  
 DB 323 LGINSTDLVLRDQFDMPSNPFGRSDLPGIDLFVSTADPEKPLVNTANTISILA 382  
 QY 379 VDPVNDKVCYSDDGAMLTFFALSETSEFAKKWPFCKRYNIEPRAEWFQOKIDYL 438  
 DB 383 VDPVPEKVSYSLDDGALLSFEMABASFDLWVFCRKHIEBRNDSYFSLKIDPT 442  
 QY 439 KDVVANFVRERAMREYEFKVRINAL-----VA 469  
 DB 443 KNSRJDIPVADRKXITREYDEFKVRINGLPDSIRRSDDAFNAEENKALKOKRESGDP 502  
 QY 470 KQKVEEGWTMODGTWMPG-----NNVRDHPGMIQVFLQSGSGALDCEG----- 514  
 DB 503 EEPKVKATW-MADGTHMPCOTMAASTREHSGKDHAIGLOVLMKRPSSDPLIGSDKVID 561  
 QY 515 -----ELPRLVYSEKRPNGVNHKAGAMNALVRSAYLTNAPYLLNLDCHYINNSK 568  
 DB 562 FSDTDLRLPMFVYVSEKRPBGYDHNNKAGAMNALVRSAYLTNAPYLLNLDCHYINNSK 621  
 QY 569 AIEAMCFMMDPLLGKVCVVOFPPRFDGIDRDRYANRVVFEIDNMGLDGIQPIY 628  
 DB 622 AVEGMCFFMMD-RGEGDICYIOPRQREBIDPSDRYANNNTVFFDGNMRALDGVQSPYV 680  
 QY 629 GTGCVFRROALYGVDAPTKKPPSRPCNCWPKMCFCCCGGNRKQKTKTKPKTEKKLLF 688  
 DB 681 GICTMRRRFPALYGFDPNPDK-----L 702  
 QY 689 FKKEENSPAYALGEIDEAAPGAENKAGIVNOQKLEKFGGSSVFTVS-----TLI 740  
 DB 703 LEKKESETEALITSDPDP-----LDVTOLPKRFGNSTLLAESPIAEFOGRPL 751  
 QY 741 ENGCTLSASP-----ASLKEAIIHVISGVEDKTDMEKIGIYASYEDIL 788  
 DB 752 ADHPAYKYGPPALRVPRLDATTVAESVSIVSCMYEDKTEWGRVGTIGSVTEVDV 811  
 QY 789 TGKQACHGRSRYICYI PKRVAFGKASPLNLSDRHLQVLRWALSGSIEIFESNHCPLVYGY 848  
 DB 812 TGRMNRNGRSRYICTRKDSFRGSAPINTDLHLQVLRWALSGSIEIFESNHCPLVYGY 869  
 QY 849 GGLKFLERSYINSIYVPMTSIDPLAYCTPAICLLTGKFTPELANNVASLWMSLFTCI 908  
 DB 870 KRLKFLORLAVLWVGVIPTSLFLIYCFPLAFSLFSGQPIVRTLISISFLVYLMITICL 929  
 QY 909 FALSILEMWSGVIDMMWNEOPVWIGVSHLPAVFOGLKVIAGVDSFTVTSK-GG 967  
 DB 930 IGLAVLEVKWSGIGLEBEMWNRNBDWMLISGTSHLVAVGVGLKVIAGIETSLITLTKSGG 989

QY 968 DDEE--FSELYTFKMTLLIPTTLNLNFIQVAVGASNAINGYESWGLPFGKLPFAFW 1025  
 DB 990 DDNEIDYADLIYVWSSIMIPPIVAMVNIATVAFIRITRYQAVPWQSKLIGCAFSEFW 1049  
 QY 1026 VIVHLYPLKGLVGRONRTPTIIVVMSLISLSLWVRIDPLADMDP 1076  
 DB 1050 VLAHLVPPAKGLMGRGKTPTIIVVWAGLITLTLWTAINP---NTGP 1096

## RESULT 14

D84741

probable cellulose synthase (imported) - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: D84741  
 R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Sher, M.; Vanken, S.E.; Unayam, L.; Tallon, J.;  
 euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventner,  
 Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84741  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-1036 <STO>

A:Cross-reference: GB:A8002093, NID:g2924781, PID:AA04910.1, GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g33100

A:Map position: 2

Query Match 36.6%; Score 2166.5; DB 2; Length 1036;  
 Best Local Similarity 44.8%; Pred. No. 1.1e-157;  
 Matches 447; Conservative 160; Mismatches 222; Indels 169; Gaps 28;

QY 178 GGCG-----KRHPPLPYADPSLPVQGRSD-----PSKDLAAYGYSVAMKE 219  
 DB 89 GGCGGPPKMGKLERRLSVMKSNKSMILRQTDGFDHNRMLFESK--GKTGIGAFWSE 146  
 QY 220 RMENMKROERHMQTNGDGGDDGDDADLPMDERAOOLSRKIPFPSSQINPYRMIIIR 279  
 DB 147 EDDTY-----DGYSKSD-----FLDKMKRLTKKQVPAILISPYRLIIYIR 189  
 QY 280 LVILGFFPHYRVNVPVNDALMLISVIGETWFMGMIIDQPKWPIERTETYLDRLSL 339  
 DB 190 LVIVFFFLWRTTNENEDAMMLWGLSIVCEIWFASFWILDLIPKLNIPNRATDIALA--- 246  
 QY 340 PDKEGQPSQLAP-----IDFVSTVDPLKEPPLVTTNTVLSLSDYDVVDKVCYSV 391  
 DB 247 HDKEQPSPSNPTGRSGLPGVDVFSVSTADPEKEPLVANTLSTLADVYPIEKLSAYIS 306  
 QY 392 DDGAMLTFFALSETSEFAKKWPFCKRYNIEPRAEWFQOKIDYLDKVAANFVRERR 451  
 DB 307 DDGAILTFPEMAEAFAVYVVPFCRKHDIERPAPDSYFISKKOPTNKKRKQDPVKORR 366  
 QY 452 AMKEEYEFKVRINALVAKAK-----VDEG-----WT 480  
 DB 367 WIKREYDEKVRINGLPQIKRAEQFMREBELKEXRIAREKNGGVLPPDGVVYKATW- 425  
 QY 481 MODGTWPG-----NNVRDHPGMIQV-----LG--QSGGIDCEGNEU--RRL 519  
 DB 426 MADGTHMPCOTWEPKPSKGDHAGIILQIMSKVPLLEVWGPNGALDFTGIDIRVPMF 485  
 QY 520 VVVSREKPGVYVNHKAGAMNALVRSAYLTNAPYLLNLDCHYINNSKAIKEAMCFMMD 579  
 DB 486 AVYSREKRPFPNNKAGAMNGMVAASAILSGAIIILNLDCHYIYNKAIKEGCFMMD 545  
 QY 580 PLGKVCVVOFPPRFDGIDRDRYANRVVFEIDNMGLDGIQPIYVGTGCVFRROAL 639  
 DB 546 -RGDRICYIOPRQREBIDPSDRYANNNTVFFDGNMRALDGLQSPVYVGTGCFRRYAL 604  
 QY 640 YGVDAPTKKPPSRPCNCWPKMCFCCCGGNRKQKTKTKPKTEKKLLFFKKEENSPAY 699

```

Db      605 YGFNPPRANEYSG-----VFQ-----EKAPAHVHTQSAQST 639
Qy      700 ALGID-BAAPGAENKAGIVNOQLEKFGQSSVFVTSTL-----LENG 743
      640 QASDLESPTQPLNDPDLG-----LPKFGNSTWF-TDTIPVAEYQGRPLADHMSYXNG 692
Qy      744 ---GTLSASG---ASLKEAIVHISCGYEEDKTMGKXIGWYGSVTEEDLTGFMHCHG 797
      693 RPPALLIPRPPLDPTVAEIAIVSCWYEDNTEGDRIGMIVGSVEDVVTGRMHRG 752
Qy      798 WRSIYCIPIKRVAFKGSAPLNTSDRLHQYLRNALGSIIEIFSNHCPALMGYGGGLTERP 857
      753 WRSYICIKRDAFPGCTAPINLTLRHLQYLRNATGSVELFFSKNNMF--AIRLKLQVR 810
Qy      858 SYINSIYVPTWSIPLLAYCTLPALCLTGKFTTPELNNVSLMFWSLFICPAT---SI 913
      811 AYLVNGIYPTFSIFLVVYCFPLALCLFSGKFTVQGLD---IHFSYLLCTIVTLTLISL 866
Qy      914 LEMNSGVGIDDMWRNEFPWYIGVSHLFAVFGQLLVIAQVDTSTFTVSK-GGDDEE- 971
      867 LEVMSGIGLEEMWNRNEQFWLIGTSAHLAAVQGLLVIAIGIEISFTLTSKASGEDED 926
Db      972 -FSELYTPKMTLLIPPTLLNLNFIQVAVGSNAINNGESWGPPLFGKLPFAFWIVHL 1030
      927 IFADLYIKMTGLFIMPLTIIIVNLVAIVIGASRTIYVITQWKGKMGIFPSLMVLTMM 986
Qy      1031 YPFLKGLVGRONRPTIYVWSIILASIFSLLWRIDP 1068
      987 YPFAKGLMGRGKVPTIYVWSGLVSIWLSLMTISP 1024

```

# RESULT 15

T02209  
 cellulose synthase (EC 2.4.1.-) catalytic chain - rice (fragment)  
 C:Species: Oryza sativa (rice)  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 11-Jun-1999  
 C:Accession: T02209  
 R:Artoli, T.; Peng, L.; Betzner, A.S.; Burr, J.; Witke, W.; Herth, W.; Camilleri, C.; H  
 Science 279, 717-720, 1998  
 A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.  
 A:Reference number: 213745; MUID:9811412; PMID:9445479  
 A:Accession: T02209  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-583 <ARI>  
 A:Cross-references: EMBL:AF030052; NID:92781432; PID:92781433  
 A:Experimental source: subsp. Japonica, strain Nipponbare  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 35.8%; Score 2118.5; DB 2; Length 583;  
 Best Local Similarity 66.2%; Pred. No. 2,4e-154;  
 Matches 398; Conservative 75; Mismatches 93; Indels 35; Gaps 8;

```

Qy      1 MEASAGLVAGSHNRNELVIRRDG--PQKPRRQNGVCOICGDDVGLAPGDPFVAC 58
      1 MAANAGMVAAGSRNRNEFMIRPDGAPPAKPKSGVNGVCOICGDTGVGSATGDPVAC 60
Db      59 NECAFPVCRDCEYERREGTGNCPOCKTRYKRLKGCORVTGDEEEDVDLDNEFMNDG 118
      61 NECAFPVCRDCEYERREGTGNCPOCKTRYKRLKGCORVTGDEEEDVDLDNEFMNDG 119
Qy      119 DSQVAESMLYGHMSYGRGDPNGAPQAFQOLNPVPLTLNGQMV-----DIPPEQHALVP 174
      120 GNGKGRPMQI-----QRGEDVDLSSSRHEQHRIPLTSGQISGEIIPDASPDHSTI-- 172
Db      175 SFMGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYIGYSVAAKERMENWKORER-MHQ 233
      173 -----RSGTSSYVDVSPVPVRIIVDPKSLNSYGVINSYDMQERVASWRNKQDKNMQ 224
Qy      234 TGN-----DGGGDDGDDADLPLMDEARQOLSRKIPLPSSQIINPYRMIIIRLVVL 283
      225 VANKTPPEARGGDMSTGNGE--DIQWVDARLPLSRIVPISNQLNLYRIYIILRLIIL 282

```

```

Qy      284 GPFPHYVMEPVNDAPALMTLSVICSIFWPMNSWILDOFPKMPRIERETLDRLSRPDK 343
      283 MFFQYVTVHPVDAYGLMLVSYICSIWFMALSWLDOFPKMYPIRKETTLDRLLARKYDR 342
Qy      344 GQPSQLAPIDFVSTVDPLKEPPLVITNTVLSILSVDPYDKVSCYVSDGAAMLTEAL 403
      343 GEPQLAPIDFVSTVDPLKEPPLITANTVLSILAVDPYDKVSCYVSDGSAAMLTEAL 402
Qy      404 SETSEPKKVVPPCKRNINIPRAPEWFOOKIDYLRKVAANFTRERANKREYEFKVR 463
      403 SETSEFARKVPPCKKNIPRAPEWFOOKIDYLRKVAANFTRERANKREYEFKVR 462
Db      464 INALVAAKQVPEEGMTMODGTPMPCNNVRDHPGMIQVPIGOSGGLDGNELPRLVYS 523
      463 INALVAAKQVPEEGMTMODGTPMPCNNVRDHPGMIQVPIGOSGGLDGNELPRLVYS 522
Qy      524 REKRPYNNHKKAGANMALVRSAYLTNAAYLNTLDCDHYINNSKAIKAMCFMMDPLG 583
      523 REKRPQGHKKAGANMALVRSAYLTNGAYLVNDCDHYINNSKALREAMCFMMDPLG 582
Qy      584 K 584
      583 R 583

```

Search completed: December 15, 2003, 13:04:59  
 Job time : 33 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:57:19 ; Search time 18 seconds

(without alignments)  
2837.278 Million cell updates/sec

Title: US-09-720-383c-10

Perfect score: 5918  
Sequence: 1 MEASAGIVAGSNHRELVI.....DEFLAKDQSPILLEECGLDCN 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	* Match Length	ID	Description
1	252	4.3	754 1	BCA1_ACEXY
2	248	4.2	874 1	BCSA_SALTI
3	246	4.2	874 1	BCSA_SALTI
4	242	4.1	739 1	BCSA_PSEFL
5	239.5	4.0	1596 1	ACST_ACEXY
6	239	4.0	872 1	BCSA_ECO57
7	238	4.0	872 1	BCSA_ECO57
8	237	4.0	756 1	BCA2_ACEXY
9	230.5	3.9	729 1	BCSA_XANAC
10	218	3.7	1518 1	BCA4_ACEXY
11	218	3.7	1518 1	BCA5_ACEXY
12	201	3.4	1550 1	ACST_ACEXY
13	194	3.3	745 1	BCA3_ACEXY
14	139.5	2.4	906 1	CBL_HUMAN
15	116	2.0	473 1	GLNA_SYNP2
16	113.5	1.9	1254 1	POLS_EEVEE
17	113	1.9	567 1	COX1_EMENT
18	113	1.9	1254 1	POLS_EEVEE
19	110	1.9	554 1	YS44_CAEEL
20	110	1.9	1912 1	CHD4_HUMAN
21	109	1.8	535 1	COX1_HANMI
22	109	1.8	896 1	CBL_MOUSE
23	108.5	1.8	357 1	G6PT_HUMAN
24	108.5	1.8	1612 1	YCDM_PARLI
25	108	1.8	441 1	YCDM_ECOLI
26	108	1.8	775 1	ICPG_HSV11
27	107	1.8	1019 1	TOP1_USDMA
28	106.5	1.8	992 1	UTRA_MICLU
29	106.5	1.8	1184 1	FBL2_HUMAN
30	106	1.8	556 1	SRW1_SCHPO
31	106	1.8	1254 1	POLS_EEVEE
32	105.5	1.8	470 1	ROCC_BACSU
33	105.5	1.8	565 1	YD08_MYCPN

34	105	1.8	873 1	COX1_ACACA	Q37370 acanthamoeb
35	105	1.8	1255 1	POLS_EEVEE	P36329 venezuelan
36	104.5	1.8	524 1	COX1_TRENT	P08741 triticum ae
37	104	1.8	537 1	COX1_SCHPO	P07657 schizosach
38	103.5	1.7	322 1	NUJM_STRPU	P15548 strongyloce
39	103.5	1.7	524 1	COX1_ORYSA	P14578 oryza sativ
40	103.5	1.7	528 1	COX1_MAIZE	P08742 zea mays (m
41	103.5	1.7	530 1	COX1_SORBI	P05502 sorghum bic
42	103.5	1.7	1120 1	KEPA_ECOLI	P77338 escherichia
43	103.5	1.7	1332 1	XKDO_BACSU	P54334 bacillus su
44	103.5	1.7	4548 1	AP0A_HUMAN	P08519 homo sapien
45	103	1.7	677 1	UBP1_XENLA	P25979 xenopus lae

## ALIGNMENTS

RESULT 1  
BCA1\_ACEXY STANDARD; PRT; 754 AA.  
AC P19449;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.  
RC STRAIN=1306-3;  
RX MEDLINE=1306-3;  
RA Wong H.C., Fear A.L., Calhoun R.D., Eichinger G.H., Mayer R.,  
RA Amikam D., Ben-Zion M., Gelfand D.H., Meade J.H., Emerick A.W.,  
RA Bruner R., Ben-Bassat A., Tal R.;  
RT "Genetic organization of the cellulose synthase operon in Acetobacter  
RT xylinum.";  
RT Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134(1990).  
RN [2]  
RP ENZYME REGULATION.  
RC STRAIN=1306-3;  
RX MEDLINE=21194571; PubMed=11297407;  
RA Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H.,  
RA Volman G., Amikam D., Ben-Zion M., Gilles-Gonzalez M.-A.;  
RT "Phosphodiesterase A1, a regulator of cellulose synthesis in  
RT Acetobacter xylinum, is a heme-based sensor.";  
RL Biochemistry 40:3420-3426(2001).  
CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose in a processive way.  
CC The thick cellulosic mats generated by this enzyme probably  
CC provide a specialized protective environment to the bacterium.  
CC -I- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl] (N) = UDP  
CC + [(1,4)-beta-D-glucosyl] (N+1).  
CC -I- COFACTOR: Magnesium.  
CC -I- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP).  
CC -I- PATHWAY: Bacterial cellulose biosynthesis.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (periplasmic).  
CC -I- INDUCTION: Cellulose is produced at a linear rate with respect to  
CC cell growth when O(2) is present.  
CC -I- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC OXKRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processive function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; M37202; AAA21884.1; -;  
 DR PIR; A43735; A43735.  
 DR InterPro; IPR003919; Cellenynth A.  
 DR InterPro; IPR001173; Glyco trans 2.  
 DR Pfam; PF00535; Glycos transf 2; 1.  
 DR PRINTS; PR01439; CELLSTHASE.  
 KM Cellulose biosynthesis; transferase; Glycosyltransferase;  
 KM Transmembrane; Inner membrane.  
 FT DOMAIN 147 240 CATALYTIC SUBDOMAIN A.  
 FT DOMAIN 317 377 CATALYTIC SUBDOMAIN B.  
 FT TRANSMEM 26 46 POTENTIAL.  
 FT TRANSMEM 47 67 POTENTIAL.  
 FT TRANSMEM 108 128 POTENTIAL.  
 FT TRANSMEM 167 187 POTENTIAL.  
 FT TRANSMEM 407 427 POTENTIAL.  
 FT TRANSMEM 430 450 POTENTIAL.  
 FT TRANSMEM 468 488 POTENTIAL.  
 FT TRANSMEM 516 536 POTENTIAL.  
 FT TRANSMEM 549 569 POTENTIAL.  
 FT ACT\_SITE 189 189 POTENTIAL.  
 FT ACT\_SITE 333 333 POTENTIAL.  
 FT BINDING 236 236 SUBSTRATE (POTENTIAL).  
 FT BINDING 238 238 SUBSTRATE (POTENTIAL).  
 SQ SEQUENCE 754 AA; 84442 MW; 805FC1FE62E2C068 CRC64;

Query Match 4.3%; Score 252; DB 1; Length 754;  
 Best Local Similarity 20.4%; Pred. No. 6.3e-11;  
 Matches 166; Conservative 89; Mismatches 217; Indels 350; Gaps 36;

QY 275 IIRIVLVGFFRRYRHPNDPFAALMSVCEIFANSMILDQ---FPRKFP----- 326  
 DB 53 LIVLVLCVLEFFIYGRKRSRTQTF-LEVISALVSLRY-LTWRLTEFLDPDWIQQGLGV 110  
 QY 327 --IERETV-LDRSLSRDEKQPSQLAP-----IDFVSTVDPPLKEPPLVTYN 371  
 DB 111 TLMAELVLYMLTSLFTOTIOPLRARPLPDNDVNDMPVDIFIPYD---EQLSLVRL 167  
 QY 372 TVLSILSVDPVVDKVCYSDSDGAAMLTFEALSETSEFAKKWVFCRKNYIEPRAPMYF 431  
 DB 168 TVLGALGIDMPDKVNVYIIDDG----- 190  
 QY 432 QOKIDYLDKVVANFVRERRAMKREYSEFKVRINALVAKAKVPEEGWTMODGTPWFGNN 491  
 DB 191 -----VRPEFQF-----AK----- 200  
 QY 492 VRDHQMIQVFLGSGSGLDCEGNLPLRLVVSREKRGVYHHKAGAMNALLVNSAVLTN 551  
 DB 201 -----DCGA-----LYIGRD-----SSHAQGLNLAIK-----RTS 228  
 QY 552 APVLLNDCHYIINNKAKEAMFMDPLIGKKVCYVOPQRFQDIDRHRYANRVNVF 611  
 DB 229 GDYILIDCDH-ITRAFLQIAMGMV---VADKIALMQPRHNYSPD----- 272  
 QY 612 FDIIMKGLDGIQGFIVYGTGCVPRQALVGYDADTKKPPSRCTNCMPKMCFCGCCFGNR 671  
 DB 273 -----PFGRLAVGY-----RTPE-----GN- 289  
 QY 672 KOKKTKPKTEKKLLFFPKKEENSPAVALGEIDEAPGANEKAGIVNQKLEKKGQS 731  
 DB 290 -----LFY-----GVITQD-----GNDPWDA 304  
 QY 732 SVFVTSITLLENGTILKSPASLKEAIVHISCGYEDKTDWKEIGVYSVTEIDITGF 791  
 DB 305 TFF-----GSCAILARKEAIESIG-----GRAVELVTEADATL 338

QY 792 KNHCHGMRSLYC-IPKRVAFKGSAPLNSDRHLQVLRMLAGSLIEIFFSNCHPLMYGGG 850  
 DB 339 RQQRMRMSTAYLRIP---VASGLATERLTLTHGQRRMRGMQIFRDNPML---GGG 391  
 QY 851 LKLEERSYINSIVPWTISIPPLAAYCTPLAICLTGKFI---TPELNNAVSLMNSLFC 907  
 DB 392 LKQGRLCYLSMNTSPFFAIPRIVILASPLALFPGQNIASP-----LA 437  
 QY 908 IFATSIENRWS---GVGIDMWRNEQFM-VIGVSSHLEFAVPGLLKVIAGVDTSFYV 963  
 DB 438 VLAVALPHEMFHSIATPAKVNKGM-RYSFWESEYETTMALFLVRVITITLMPFSKGFNT 496  
 QY 964 SKGG--DDEPSESLYFFKKTITLIPPTITLLNPLFGVAVGSNAINNGESGPIFG-K 1019  
 DB 497 EKGGLVEEFPD-----LGATYPNIIIPAGIMTGLLGLFE 532  
 QY 1020 LPEAFVVIHLVPLFKGLVGRNRPPTIVYS-----ILLASI 1058  
 DB 533 LTFHF-----NQLAGIAKRAYLNCINAMISLITLAI 566

RESULT 2  
 ID BCSA\_SALTI STANDARD; PRT; 874 AA.  
 AC 082331;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12).  
 GN BCSA OR STY4181 OR T3898.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxId=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagsis K.,  
 RA Krogan A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18";  
 RL Nature 413:848-852(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ty2 / ATCC 700931;  
 RX MEDLINE=22531367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 RA Birlind V., Kodoyianni V., Schwartz D.C., Blatcher F.R.;  
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
 RT and CT18";  
 RL J. Bacteriol. 185:2330-2337(2003).  
 CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
 CC an extracellular component for mechanical and chemical protection  
 CC at the onset of the stationary phase, when the cells exhibit  
 CC multicellular behavior (rod morphology). Co-expression of  
 CC cellulose and thin aggregative fimbriae leads to a hydrophobic  
 CC network with tightly packed cells embedded in a highly inert  
 CC matrix (by similarity).  
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP  
 CC + ((1,4)-beta-D-glucosyl)(N+1).  
 CC -1- COFACTOR: Magnesium (By similarity).  
 CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (c-di-GMP) (By similarity).  
 CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

```

CC (Potential).
CC -!- DOMAIN: There are two conserved domains in the globular part of
CC the protein: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRM motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL, AL627281, CAD08006.1, -.
DR EMBL, AE016847, AA071373.1, -.
DR InterPro, IPR003919, CellSynth_A.
DR InterPro, IPR001173, Glyco_transf_2.
DR Pfam, PF00535, Glycosyltransferase_2, 1.
DR PRINTS, PR01439, CELL5NTHASRA.
KM Cellulose biosynthesis; Transferase; Glycosyltransferase;
KM Transmembrane; Inner membrane; Complete proteome.
FT DOMAIN 271 364
FT TRANSMEM 441 501 CATALYTIC SUBDOMAIN A.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 525 545 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 592 612 POTENTIAL.
FT TRANSMEM 634 654 POTENTIAL.
FT TRANSMEM 668 688 POTENTIAL.
FT TRANSMEM 833 853 POTENTIAL.
FT ACT_SITE 313 313 POTENTIAL.
FT ACT_SITE 457 457 POTENTIAL.
FT BINDING 360 360 SUBSTRATE (POTENTIAL).
FT BINDING 362 362 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 874 AA, 100020 MW, 339509 CD05F1990 CRC64;

Query Match 4.2%; Score 248; DB 1; Length 874;
Best Local Similarity 18.0%; Pred. No. 1.6e-10;
Matches 152; Conservative 100; Mismatches 198; Indels 394; Gaps 32;

258 LSRKLPSSQINPRMIIIRLVLGFFHYRVHNPV--DAFAL-----W 302
189 LVRRMRGRPSAL-----MLIVLSLVSCRYIMRWYSTLWMDPVSILVCGILILFAETVAM 244
303 LLSVLCIEIFWAMSWILDQPPKMPFIERETVLDLSRPFKEQPSGLAIDPFVSIVDVL 362
245 IYLVLG--FYQVWVPLRQRP--VPLPKEM-----SQMPTVIDFVPTVY-- 283
363 KKPPIVTTNTVLSISVDYPRVDKSCYVDDGAAMLTPEALSETSEFAKMWPFCKRYNI 422
284 -EDLVVVKRTIYASIGIDWPVKDKNTIWIIDGG----- 315
423 EPRAPRWYFOQKIDVLDKDVAAANFVRERAPMKREYEFKRVINALVAKQVPEEGWTNQ 482
316 -----RESFRQF----- 322
483 DGTWPNGNNVRDHPGMIQVFLGSGGLDGEENLPRLLVVSREKRPVHHKKAGAMNL 542
323 -----ARRVGVH-----YIAR-----ATHEHAKAGINNA 347
543 VRVSAVLTAPYLLNDCDHYIINNSKAIEAMCFMMDPLGKVCVQVQPFQFQDID-- 599
348 LKHA-----KGFPAIRGDCDHVPRRSLQMTMGWFLKE-----KQLAMWQPRHHFSPDPE 399

```

```

QY 600 ----RHDRYANRVNVPFDINKKGLDGIQGIYVGHGCVFRRQALYGYDAPKTKPPSRTC 655
DB 400 RNLGRFRTPRBEGLTFYGLVODGND----- 424
QY 656 NCWPMWCCGCCCGFNKKOKKTKTPKTEKKLLFFKKEENQSPAYALGIDEAIPAAGANEK 715
DB 425 -MMDATPFCGSC----- 435
QY 716 AGIVNQGLEKKFGQSSVFVSTLLENGTLKASAPASLBALHIVISGYEDKTDWKE 775
DB 436 -AVIRKRLD-----E 445
QY 776 IGMVYGSVEDILTFPMHCHGMRSIYC-IPKRYAFKGSAPLNLSDRLHOVLALGSI 833
DB 446 IGGIVETVETDAPTSLSLRGRTSAYMRAPQSA--GLATESLSAHIGQIRIMARQV 502
QY 834 EIFFSNHCPVWYGGGLKFLERSYINSIYVPMWISPLAVCTLPALCLTGKPTPEL 893
DB 503 QIFRLDN-PL--FKGKLKLRCLCYLNAHHPISGIRLLFLTPALFLHAYI----- 554
QY 894 NNVASLMFMSLFICIFA-----TSILEMRSQVIGIDDMWRNEQF-----WVGVSQHL 942
DB 555 -----IYAPALMIALFVLPWVHASLTNSKIQGKRYHSFM-SEIYETVLAWYIAPT--- 605
QY 943 FAVPGGLKVIAGVDTSTVTSKSGDDEEPESELYTFKKTILLI--PPTLLILNPIGVA 1000
DB 606 -----LVALLNPKKKNVTKAGGLVEE-----KVDVWISRPYFLVLLNLGVA 652
QY 1001 GVSNAINNGYSWGPFLPKLFFAFVVIYHLYPLKGLGVRONRPTVIYVSIILASIFS 1060
DB 653 GV-----WRYLYGP-----ENETLITVIV-----S 671
QY 1061 LLMV 1064
DB 672 LWMV 675

RESULT 3
BCSA_SALTY
ID BCSA_SALTY STANDARD; PRT; 874 AA.
AC Q93IN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR STM3619.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14028;
RX MEDLINE=21160181; PubMed=11260463;
RA Zogaj X., Nitez M., Rohde M., Bokranz W., Roemling U.;
RT "The multicellular morphotypes of Salmonella typhimurium and
RT Escherichia coli produce cellulose as the second component of the
RT extracellular matrix.";
RL Mol. Microbiol. 39:1452-1463(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21927388; PubMed=11929533;
RA Solano C., Garcia B., Valle J., Beraain C., Ghigo J.-M., Gamazo C.,
RA Lasa I.;
RT "Genetic analysis of Salmonella enteritidis biofilm formation:
RT critical role for cellulose.";
RL Mol. Microbiol. 43:793-808(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McCelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

```

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
 CC an extracellular component for mechanical and chemical protection  
 CC at the onset of the stationary phase, when the cells exhibit  
 CC multicellular behavior (rod morphology). Co-expression of  
 CC cellulose and thin aggregative fimbriae leads to a hydrophobic  
 CC network with tightly packed cells embedded in a highly inert  
 CC matrix.  
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl](N) = UDP  
 CC + [(1,4)-beta-D-glucosyl](N+1).  
 CC -1- COFACTOR: Magnesium (By similarity).  
 CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (c-di-GMP) (By similarity).  
 CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Periplasmic).  
 CC -1- DOMAIN: There are two conserved domains in the globular part of  
 CC the protein: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXXW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AJ151770; CAC44015.1; -;  
 CC EMBL: AE008867; AAL22479.1; -;  
 CC EMBL: AJ151548; CAC86199.1; -;  
 CC StryGene: SG27272; bcsA.  
 DR InterPro: IPR003919; CellSynth\_A.  
 DR InterPro: IPR005150; Cellulose\_synth.  
 DR InterPro: IPR001173; Glyco\_trans\_2.  
 DR Pfam: PF03552; Cellulose\_synth\_1.  
 DR Pfam: PF00535; Glycos\_transf\_2; 1.  
 DR PRINTS: PR01439; CELLSTHASEA.  
 KM Cellulose biosynthesis; transferase; glycosyltransferase;  
 KM Transmembrane; Inner membrane; Complete proteome.  
 FT DOMAIN 271 364  
 FT TRANSMEM 441 501 CATALYTIC SUBDOMAIN A.  
 FT TRANSMEM 30 50 CATALYTIC SUBDOMAIN B.  
 FT TRANSMEM 151 171 POTENTIAL.  
 FT TRANSMEM 173 193 POTENTIAL.  
 FT TRANSMEM 230 250 POTENTIAL.  
 FT TRANSMEM 525 545 POTENTIAL.  
 FT TRANSMEM 547 567 POTENTIAL.  
 FT TRANSMEM 592 612 POTENTIAL.  
 FT TRANSMEM 634 654 POTENTIAL.  
 FT TRANSMEM 668 688 POTENTIAL.  
 FT TRANSMEM 833 853 POTENTIAL.  
 FT ACT\_SITE 313 313 POTENTIAL.  
 FT ACT\_SITE 457 457 POTENTIAL.  
 FT BINDING 360 360 SUBSTRATE (POTENTIAL).  
 FT BINDING 362 362 SUBSTRATE (POTENTIAL).  
 SQ SEQUENCE 874 AA; 100043 MW; 4C9421B58606310A CRC64;

Query Match 4.2%; Score 246; DB 1; Length 874;  
 Best Local Similarity 18.1%; Pired. No. 2.2e-10;

Matches 153; Conservative 99; Mismatches 199; Indels 394; Gaps 32;  
 QY 258 LSRKIPDPSSQINPYRMIIIRLVYGFPPHYRVHNPVN--DAFAL-----W 302  
 DB 189 LYRMRGRFSAAL---MLIVSLTVSCRYIMWRYSSTLNMDPVSGLGILLFAETVAM 244  
 QY 303 LISVCEIWFAMSTLDDQPPKMFPIERTYIDRLSRFDKSGQSLAPIDPFYSTVDP 362  
 DB 245 IYLVIG--YFQVWPLNRQP--VPLPKEM-----SQWPTVDIVFPTVN-- 283  
 QY 363 KEPLVTNTVSLTSLVDYVDKVSQVSDGAMLTREALSTSEFAKKVPPCKRNI 422  
 DB 284 -EDLVNKKTVIASIGIDWPDKNTWILDQGR----- 316  
 QY 423 EPRADENYFOQKIDYLDKQVANAFFRRBRAMKREYEFKVRINALVAKQVPEEGMTQ 482  
 DB 317 -----ESFR----- 320  
 QY 483 DGTMPGNVNDHPGMIVFLIGQSGGLDCGNEELPRLVYVEREKPGRYNNHKKAGANAL 542  
 DB 321 -----HFAHVG-----VHIARTT---HEHAKAGININA 347  
 QY 543 VRVSAVLINAPVILNLDGDHYINNSKAIKEMCFMMDLKKVCYVQFPQPDID--- 599  
 DB 348 LKHA---KGEFVAIFDDCHVPTRSFLQMTGMFLKE---KQLAMQTPHHFSPDPPE 399  
 QY 600 ---RHDRYANRVNVPFDINMKGLDGIQPIYVGTGCVFRRQALGYDAPRTKPPSRTC 655  
 DB 400 RNLGRFRKTPNGCTLPYGLVQDGD----- 424  
 QY 656 NCPFRMCFCCCGFNGRKKQTKTKPTEKKLLFFKEENQSPAYALGIDEAPGAENEK 715  
 DB 425 -MMDATFCGSC----- 435  
 QY 716 AGIVNQQLKEKFGGSSVFTVSTLLENGTLKSASPAALKEAHNVISGVEDTKDWKE 775  
 DB 436 -AVIRKPLD-----E 445  
 QY 776 IGMV-YGSYTEDILTFKMKHCHGMSIYC-IPKRVAFKGSAPLNDRLHQLVRLALGSI 833  
 DB 446 IGGIAVEVYTBDAHSLRLHGRGYSAYMRIPQA--GLATESASAHIGQIRARQMV 502  
 QY 834 EIFFSNHCPLMWYGGGLKFLERFSYINSIYVPTWSTIPLAYCTLPALCLLTGKFTPEL 893  
 DB 503 QIFRLDN-PL---FEGKGLKQRLCYLNAMFHLGIPRLIFLTPALFLHLHAVY---- 554  
 QY 894 NNVASLMFMSLFCIFA-----TSILEMRNGVGIDDMWRREOF-----WIGGVSSHL 942  
 DB 555 ---IYAPALMTALFVLIPIHVVHASTNSKIQGRHSFW-SEIETVLAWTIADPT--- 605  
 QY 943 FAVFQGLKVIAGVDTSTFTVTSKGGDDEFESELYTFKWTLLI--PTTLLLNFGVVA 1000  
 DB 606 -----LVALINPHKGRVNTAKGLVE-----KYDVWISRPYIFVLVNLGLVAA 652  
 QY 1001 GVSNAINNGYESWGLFGCLFFAFVIVHLVPLKGLVGRONRPTIIVISILLASIFS 1060  
 DB 653 GV-----WRYYYGP-----EBETLTIVY-----S 671  
 QY 1061 LLNV 1064  
 DB 672 LVNV 675  
 RESULT 4  
 BCSA\_PSEFL STANDARD; PRT; 739 AA.  
 AC P58931; Q8RSZ1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12).  
 GN BCSA OR WSSB.  
 OS Pseudomonas fluorescens.

CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 CC Pseudomonadaceae; Pseudomonas.  
 CC NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SBW25;  
 RA MEDLINE=22013850; PubMed=12019221;  
 RA Spiers A.J., Kahn S.G., Bohannon J., Traviano M., Rainey P.B.;  
 RT "Adaptive divergence in experimental populations of *Pseudomonas*  
 RT *fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader  
 RT fitness.";  
 RL Genetics 161:33-46(2002).  
 CC  
 CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
 CC an extracellular component responsible for the structural  
 CC integrity and rigidity of self-supporting mats characteristic of  
 CC the "wrinkly spreader" phenotype.  
 CC -I- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl) (N) = UDP  
 CC + ((1,4)-beta-D-glucosyl) (N+1).  
 CC -I- COFACTOR: Magnesium (By similarity).  
 CC -I- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (C-di-GMP) (By similarity).  
 CC -I- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -I- DOMAIN: There are two conserved domains in the globular part of  
 CC the protein: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXXRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 CC  
 DR EMBL: AY074776; AAL71842.1; .  
 DR InterPro: IPR001173; Glyco\_transf\_2.  
 DR Pfam: PF00535; Glycos\_transf\_2; 1.  
 KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 KW Transmembrane; Inner membrane.  
 FT DOMAIN 157 250  
 FT DOMAIN 327 387  
 FT TRANSMEM 36 55  
 FT TRANSMEM 59 76  
 FT TRANSMEM 83 101  
 FT TRANSMEM 116 138  
 FT TRANSMEM 417 436  
 FT TRANSMEM 440 462  
 FT TRANSMEM 524 546  
 FT TRANSMEM 551 573  
 FT ACT SITE 199 199  
 FT ACT SITE 343 343  
 FT BINDING 246 246  
 FT BINDING 248 248  
 FT SEQUENCE 739 AA; 82165 MW; 2B962E83854B23BB CRC64;  
 SQ  
 Query Match 4.1%; Score 242; DB 1; Length 739;  
 Best Local Similarity 19.1%; Pred. No. 3.5e-10;  
 Matches 155; Conservative 86; Mismatches 207; Indels 364; Gaps 33;  
 QY 260 RKIPVSSQINRYRMIIIRLVV-----LGF-----FFHYRVHHPVN 296  
 DB 77 RKIP-----GRLALIALVYLSVLSARYMFRRLNSTLGFETWDMFPGYGLV---- 123  
 QY 297 DAFALMLISVCEIWFAMSWIUDQPPKPFIERETYLDRLSIRPKGQPSQLADIPDFV 356

DB 124 -AAEFYALIVIFGVQTAMPRLRTPYMKTE-----PESMTVDVFI 165  
 QY 357 STDVPLEKPLVTNTVLSILSVDPYDKVSCYVSDGAAMLTREALSETSEFAKWP 416  
 DB 166 PTYN---EALSIVKLTIFAQAMDMPKDKLRVHVLDDG-----RRDDFRE----F 208  
 QY 417 CKRNIEPRAPEWIFQOKIDYLDKQVNAFPRERAKRKEVEFKRINMLVAQAQVPE 476  
 DB 209 CR-----KGVNYIR----- 218  
 QY 477 EGMTWQDCTPWRGNVRDHPGMIQVPLGGSGGLDEGNELPRLVYVSREKRQYHHKKA 536  
 DB 219 -----RDN-----NFAKA 227  
 QY 537 GAMNALVRSVAVLTNAPYLNTLDCDHYINNSKAKEAMCFMMDPLGKVCYVQPGRF 595  
 DB 228 GNLNEALKV-----TDGEIALFDADHVPPTSFLQVSLGWFLKDB-----KLAMLTGPHPEF 279  
 QY 596 -----DGLDHDHRYARNRVVFPDINKGDLGIGPIYVGTGCVFRQALYGYDAPTKK 649  
 DB 280 SPDPPEKRLDTFRAVPNEGELFVGLVODGND----- 310  
 QY 650 PPSRTCMQWRKWCRCGCCCFGRKQKTKTKPTEKKKLLFKKENQSPAYALGIDEAP 709  
 DB 311 -----LMAATPFC----- 318  
 QY 710 GAENEKAGIVNQKLEKKFGQSVFVSTLLENGTLKSPASLLEKAIHVISGVEDK 769  
 DB 319 -----GSCAVIRREPLLEIGV-----AVG----- 338  
 QY 770 TDWKEIGIWTYGVSTEDILTFGRKHCQWRSY-CIPRVAFKSAPLNSDRLLHQLVR 828  
 DB 339 -----TYTEAHNTALKNRLGYNVAYLAIPQAA-----GLATESRHHINQRIWR 383  
 QY 829 ALGSIETIFSNHCLMVGGLGKLEKRFYSYINSIVYPTWISPIILACTPLAICLTCKF 888  
 DB 384 ARGMAQIFRTN-PL--LGGLKMGQRTICANAMQHFYGLPRLVFLTAPLAYLIFG-- 437  
 QY 889 ITPELNNVSAIWMFLFI-CIFATSIEMRWGVGIDDMWNEOF-----WVGVSSTL 942  
 DB 438 --AIFHASALMIVAYVLPVHSSLTNSRIGQRFHSFW-NEVYEVTLAMYI-----L 488  
 QY 943 FAYFOGLKVIAGVDTSTVTSKCG-DDEESELVTFKMTLLIPPTLLINPFGVAG 1001  
 DB 489 PPVVALVNPVRAQ---GFTVVDKGGIIDKQF---FDW-KIARPYLVLLAVNLIQLGFG 539  
 QY 1002 VSNAINNGYESMGPLFGKLPFAF---WVIYHL 1030  
 DB 540 IHOIL-----WGDASTAVVAIVLWTLVNL 565  
 RESULT 5  
 ACS2\_ACEXY STANDARD; PRT; 1596 AA.  
 AC Q59167;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cellulose synthase 2 [includes: Cellulose synthase catalytic subunit  
 DE (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose  
 DE synthase 2 regulatory domain)].  
 GN ACSA11.  
 OS Acetobacter xylinus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
 OC Acetobacteraceae; Gluconacetobacter.  
 OX NCBI\_TaxID=28448;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 23769;  
 RX MEDLINE=95394846; PubMed=7665515;  
 RA Saxena I.M., Brown R.M. Jr.;  
 RT "Identification of a second cellulose synthase gene (acsA11) in

```

RT Acetobacter xylinum".
RU Bacteriol. 177:5276-5283(1995).
CC -I. CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -I. SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -I. DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processive glycosyl
CC function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -I. MISCELLANEOUS: It is not essential for cellulose production in
CC this strain.
CC -I. SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC GLYCOSYLTRANSFERASE FAMILY 2.
CC -I. SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSE/BCSB
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U15957; AAA85264.1; -.
DR PIR: T31338; T31338.
DR InterPro: IPR003919; Cellsynth_A.
DR InterPro: IPR003920; Cellsynth_B.
DR InterPro: IPR001173; Glyco_trans_2.
DR Pfam: PF03170; Bcsb; 1.
DR Pfam: PF00535; Glycos_transf_2; 1.
DR PRINTS: PR01439; CELLSTHASE.
DR PRINTS: PR01440; CELLSTHASEB.
KW Cellulose biosynthesis; transferase; glycosyltransferase;
KW Transmembrane; Inner membrane.
KW -----
FT DOMAIN 1 749 CATALYTIC.
FT DOMAIN 750 1596 CYCLIC DI-GMP BINDING (POTENTIAL).
FT DOMAIN 145 238 CATALYTIC SUBDOMAIN A.
FT DOMAIN 315 375 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 396 416 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT TRANSMEM 505 525 POTENTIAL.
FT TRANSMEM 544 564 POTENTIAL.
FT TRANSMEM 1553 1573 POTENTIAL.
FT ACT SITE 187 187 POTENTIAL.
FT ACT SITE 331 331 POTENTIAL.
FT BINDING 234 234 SUBSTRATE (POTENTIAL).
FT BINDING 236 236 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 1596 AA; 175799 MW; D98A6F6259E1F3CE CRC64;
Query Match 4.0%; Score 239.5; DB 1; Length 1596;
Best Local Similarity 18.2%; Pred. No. 1.6e-09;
Matches 145; Conservative 91; Mismatches 228; Indels 331; Gaps 26;
QY 274 MIIIRLVLGFFPHVKNHPVNDAPALMILIS-----VICETWPMNSILDOFPKMP 326
DB 81 LTVSLRYMVRILTTEIHSPLQNALSLILVAELVALTLTCLSYFQMSWPLDRKP--LP 138
QY 327 IREETVYDRSLRPNKSGQSQLAFIDFVSTVDPLKEPPLVNTNTLTSLSDVYPDVKY 386
DB 139 LPADT-----TDMPVVVVYVPSN---EELSVRSYVLGALADWPDKL 180
QY 387 SCVYSDGAAMLTFEALSETSEFAKAVVPCKRYNIPRAPEWYFOOKIDYLKDKVAANF 446
DB 181 NVYILDG----- 188

```

```

QY 447 VERBAMKREYEFKVRINALVAKAKVPEEGMTWQDGPMPGNVNDHPGMIGVFLGDS 506
DB 189 -----RKRSFAFAMEGA-----GYIRIQ----- 209
QY 507 GGLDCEGNELPRLVYVSEKRGYNNHKKAGAMNLYVSAVLTNAPYLMLDCDHYINN 566
DB 210 -----NNHAKAGNLNALKV-----TEGEYVIFDCGH-IPT 240
QY 567 SKAIEAMCEPM-DPLGKVCYVQPPQFPGIDRHR--YANRV-----VFEDINMG 618
DB 241 RGFPLKKTIGMMWADP-----KLLAQTPHHFYSPPDFORNLATGVNVPPEGNMYGLVQDG 296
QY 619 LDGIGPIYVGNCGVFRQALYGVDAKPKTKPRSTCNCPWKCCECCCFGRKQKTKTK 678
DB 297 NDFMDATPFCGSCAARSAVIGIG----- 322
QY 679 PTEKKKLLFFKKEENOSPAYALGIDBAFGENEKAGIVNQLEKKFGQSVFVST 738
DB 323 -----FATET 327
QY 739 LIENGGLTKASPASLKEALHVISGYEDKTDWKEIGWYGVTEIDLTGFKMCHGM 798
DB 328 -----VTEDAHITALKMQREGM 343
QY 799 RSIYCIKPKVAKGSAPLNLSRNLQVLRMALGSIIEFNSNCPMYGGGLKFLERS 858
DB 344 HRAVL--KQPLAAGSTERLMIHIDQVRMAQMIIQIRLND--PL--LGSGLRQQRIC 397
QY 859 YINSIYVPTSTIPLAYCTLPALICLTGKFTPELNNVNASLTFMSL-----FTICPAT 911
DB 398 YLSAMSHFLFALPRLVFLASPLAFLEQG-----NIIAASFPALVYAFPHVFSIGTL 451
QY 912 SLENRMWQSVGIDDMRNEQFM-VIGVSSHLFAVPOGLKVIAGVDSFTYTSKG-DD 969
DB 452 SVEGGRM-----RYSFWEIYETTLAFLVRVITLTLNDRKGFENVTDKXGLIQ 501
QY 970 EEFSELYTFKWTLLIPPTLLLNFIQVAVASNAINNCGSWKBPJCKLFFAFVVIYH 1029
DB 502 SEYFPLD-----NAYVPVITLAVITALLVRIQIGMWE-YHRLALQSFALNTLWAVS 554
QY 1030 LYPFLKGL-VGRQNR 1043
DB 555 LIIVLASINAGRETR 569
RESULT 6
BCSA_ECO57 STANDARD; PRT; 872 AA.
ID BCSA_ECO57
AC 08X5T7;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR 24948 OR EC54413.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck B.V., Davis N.W., Lim A., Dimalanta E.T., Poromousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Weich R.A., Blatterer F.R.;
RT Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;

```

RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.",  
 RL DNA Res. 8:11-22(2001).  
 CC - FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5',-diphosphate glucose to cellulose, which is produced as  
 CC an extracellular component for mechanical and chemical protection  
 CC at the onset of the stationary phase, when the cells exhibit  
 CC multicellular behavior (rod morphology). Co-expression of  
 CC cellulose and thin aggregative fimbriae leads to a hydrophobic  
 CC network with tightly packed cells embedded in a highly inert  
 CC matrix (By similarity).  
 CC - CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl) (N) = UDP  
 CC + ((1,4)-beta-D-glucosyl) (N+1).  
 CC - COFACTOR: Magnesium (By similarity).  
 CC - ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (c-di-GMP) (By similarity).  
 CC - PATHWAY: Bacterial cellulose biosynthesis.  
 CC - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC - DOMAIN: There are two conserved domains in the globular part of  
 CC the protein: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXXRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC - SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 CC -----  
 CC EMBL, AB005579; AAC58675.1; ALT\_INT.  
 CC EMBL, AB002565; BAB37836.1; ALT\_INT.  
 CC DR InterPro: IPR003193; Glyco trans. A.  
 CC DR InterPro: IPR001173; Glyco trans. 2.  
 CC DR Pfam: PF00535; Glycos transf. 2; 1.  
 CC DR PRINTS: PR01439; CELLSTHASEA.  
 CC DR Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 CC Transmembrane; Inner membrane; Complete proteome.  
 CC KW DOMAIN 271 364 CATALYTIC SUBDOMAIN A.  
 CC FT DOMAIN 441 501 CATALYTIC SUBDOMAIN B.  
 CC FT TRANSMEM 30 50 POTENTIAL.  
 CC FT TRANSMEM 151 171 POTENTIAL.  
 CC FT TRANSMEM 173 193 POTENTIAL.  
 CC FT TRANSMEM 230 250 POTENTIAL.  
 CC FT TRANSMEM 525 545 POTENTIAL.  
 CC FT TRANSMEM 547 567 POTENTIAL.  
 CC FT TRANSMEM 592 612 POTENTIAL.  
 CC FT TRANSMEM 640 660 POTENTIAL.  
 CC FT TRANSMEM 668 688 POTENTIAL.  
 CC FT TRANSMEM 833 853 POTENTIAL.  
 CC FT ACT\_SITE 313 313 POTENTIAL.  
 CC FT ACT\_SITE 457 457 POTENTIAL.  
 CC FT BINDING 360 360 SUBSTRATE (POTENTIAL).  
 CC FT BINDING 362 362 SUBSTRATE (POTENTIAL).  
 CC SQ SEQUENCE 872 AA; 99710 MW; F3F1A24A2B713FBA CRC64;  
 CC  
 CC Query Match 4.0%; Score 239; DB 1; Length 872;  
 CC Best Local Similarity 17.9%; Pred. No. 7.5e-10;  
 CC Matches 147; Conservative 96; Mismatches 196; Indels 380; Gaps 33;  
 CC 274 MIIIRLVVIGFFPHVMPVNDAPALM--LISVCEI-----WFRAMWILDQFP 322

Db 201 MIVISLTVSCTRIWRYSTLN-----WDDPSLVGGLILPAETVAVIV--VIGYRQ 253  
 QY 323 KMFPIREFYLDRLSRFDKEGQSQLADIDFVSTVDPLKEPPLVTTVLSISVDY 382  
 Db 254 VVWPLNRQP---VPLPKDMSLWPS---VDIFVPTYN---EDLVNKKVTIVASLGIDWP 302  
 QY 383 VDKVSCYVDDGAAALTFPALSTSEFAKKWPFCKRYNIEPRAPEWFOQKIDYKDY 442  
 Db 303 KDLNLTWILDDG----- 315  
 QY 443 AANFVBERAMKREVEEFKVRINALVAKAKVPEEGMTQDDPWPQGNVVRDPMIOYF 502  
 Db 316 -----REFRQF-----AQN----- 326  
 QY 503 LQSGGLDCEGNELPLVYVSRKRGYNHKKAGAMNLVRSAYLTAPYLNLDCDH 562  
 Db 327 -----GVKTIART---HEHAKAGINNALKYA---KGFVSIIDCH 363  
 QY 563 YINNSALKEAMCFMMDPLGKVCYVQFPQFQID-----RDRYANRVVFPDIN 615  
 Db 364 VPTRSPLQMTVGWFLKE---KQLAMQTPHHFSPDPERNLGRFRKTPNEGTLFYGLV 419  
 QY 616 MKGLDGIQPIYVGTGCVFRQALYGYDAPKTKKPSRTCNCPKCFCCCGRNKKKK 675  
 Db 420 QDQND-----WMDATPFGSC----- 435  
 QY 676 TTPKTEKKLLPFKKENQSPAYALGEIDBAPEAKENKAGIVNOCKLEKKRQGSVPV 735  
 Db 436 -----AVIRKPLD----- 444  
 QY 736 TSTLLNGTGLKASAPASLKEAIVHISCGEDKTDWKEIGMI-YGVYTEDILTFKRM 794  
 Db 445 -----ELGGIAVEITVEDAHTSLRLH 465  
 QY 795 CHQWRSLYC-IPKRAVFKGAPLNSDLRHOVLRLALGSIETFFSHCHLWYGGGLYF 853  
 Db 466 RRYGTAYRIRIPAA--GLATESLSAHIGQIRMRMGWQVIFRLD-PL--TGGLKE 518  
 QY 854 LERFSYINSIVYWTISIPLAYCTLPALCLTGKFTPELNVANVASFMSFCIFPA--- 910  
 Db 519 AQLCLVNMAMFHLISIPPLIFITAPLALHLHAYI-----IYALMLALVLRH 569  
 QY 911 ---TSLERKMSGVGIDWNRNQF---WVIGVSHLFAVGQLKVIAGVDTSPY 962  
 Db 570 MIHASTNSKIQCKYRHSFW-SRIYETVLAWYIAPPT-----LVALLNHKKGFNV 619  
 QY 963 TSKCG-DDEEFSSELYFKWTTLLIPPTLLNFGVAVGVSNAINNGESWGLFGKLF 1021  
 Db 620 TAGGLAVEEYVD---W-VISRPYIFVLNLVGVAVGI-----WRYFYG--- 660  
 QY 1022 FAFWVIVHLPPFLKGLVGRQRNRTPTIVVSIILASIFSLWV 1064  
 Db 661 -----PPT-----ENLTIVVSMWV 675  
 CC  
 CC RESULT 7  
 CC BCSA\_ECOLI STANDARD; PRT: 872 AA.  
 CC ID P37653; P37654; P76712; P76713; Q8R8S7;  
 CC DT 01-OCT-1994 (Rel. 30, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12).  
 CC GN BCSA OR B3533.  
 CC OS *Escherichia coli*.  
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC OC Enterobacteriaceae; *Escherichia*.  
 CC OX NCBI\_TaxID=562;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=K12 / MG1655;  
 CC RX MEDLINE=94316500; PubMed=8041620;





QY 1022 FAFWTVHLYPELKLGVGRNRTPIVWISLILASIFSILMV 1064  
 Db 661 -----PPT-----EMLTIVVSSMWV 675

RESULT 8  
 BCAA2\_ACEXY  
 ID\_BCAA2\_ACEXY STANDARD; PRT; 756 AA.  
 AC 082859;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
 GN BCSA.  
 OS Acetobacter xylinus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
 OC Acetobacteraceae; Gluconacetobacter.  
 OX NCBI\_Taxid=28448;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BPR 2001;  
 RX MEDLINE=98296257; PubMed=9630539;  
 RA Nakai T., Moriya A., Tonouchi N., Tsuchida T., Yoshinaga F.,  
 RA Horinouchi S., Sone Y., Mori H., Sakai F., Hayashi T.;  
 RT "Control of expression by the cellulose synthase (bcsa) promoter  
 RT region from Acetobacter xylinum BPR 2001.";  
 RL Gene 213:93-100(1998).  
 CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5'-diphosphate glucose to cellulose. The thick cellulosic  
 CC mats generated by this enzyme probably provide a specialized  
 CC protective environment to the bacterium (By similarity).  
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP  
 CC + ((1,4)-beta-D-glucosyl)(N+1).  
 CC -1- COFACTOR: Magnesium (By similarity).  
 CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (c-di-GMP).  
 CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- DOMAIN: There are two conserved domains in the globular part of  
 CC the catalytic subunit: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXKRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, AB010645; BAA31463.1; -;  
 CC InterPro: IPR003919; Cellulynth A.  
 CC InterPro: IPR001173; Glyco\_transf\_2.  
 CC Pfam: PF00535; Glyco\_transf\_2; 1.  
 CC PRINTS: PR01439; CELLSNTHASEA.  
 CC Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 CC Transmembrane; Inner membrane.  
 CC FT DOMAIN 147 242 CATALYTIC SUBDOMAIN A.  
 CC FT DOMAIN 319 379 CATALYTIC SUBDOMAIN B.  
 CC FT TRANSMEM 27 47 POTENTIAL.  
 CC FT TRANSMEM 49 69 POTENTIAL.  
 CC FT TRANSMEM 106 126 POTENTIAL.  
 CC FT TRANSMEM 167 187 POTENTIAL.  
 CC FT TRANSMEM 409 429 POTENTIAL.  
 CC FT TRANSMEM 432 452 POTENTIAL.

FT TRANSMEM 470 490 POTENTIAL.  
 FT TRANSMEM 517 537 POTENTIAL.  
 FT TRANSMEM 551 571 POTENTIAL.  
 FT ACT\_SITE 189 189 POTENTIAL.  
 FT ACT\_SITE 335 335 POTENTIAL.  
 FT BINDING 238 238 SUBSTRATE (POTENTIAL).  
 FT BINDING 240 240 SUBSTRATE (POTENTIAL).  
 SQ SEQUENCE 756 AA; 84561 MW; 6954F39A25E73B0A CRC64;

Query Match 4.0%; Score 237; DB 1; Length 756;  
 Best Local Similarity 19.2%; Pred. No. 8.6e-10;  
 Matches 155; Conservative 96; Mismatches 247; Indels 308; Gaps 31;

QY 269 INPRTMIILNVLGFEFHVRVHVPNDAPALMLISVCEIWMAMGILDQ---FPRKF 325  
 Db 47 LNNBQLIVAACVAVIFVVGKGRSRTQIF-LEVLSSALVRLY-LTRRLTETLDFNTWI 104  
 QY 326 P-----IERETV-LDRLSLRFKQGPQSLAP-----IDFVSTVDPLEK 365  
 Db 105 QGILGVLLMABELVALVYLFLSYFQTIQPLHRAILPLPDNDVDPYDIFIPYD---EQ 161  
 QY 366 PLVTTNTVLSLSDYDPVDKVCYVSDGAAMLTFEALSTSEPAKKWPFCKRYNIBPR 425  
 Db 162 LSIYRLTVLGLGIDWPDKNVYILDDG----- 190  
 QY 426 APENYFQKKIDLYDKVAANVRRBRAMKREYSEPKRYNALVAKQKVPSEGMTQDGT 485  
 Db 191 -----VRPEEOPADCGALYI----- 207  
 QY 486 PMPGNVVDHPGMIQVFLGSGGLDCEGNEPLRLVYVSREKPGNHHKKAGANALVRV 545  
 Db 208 -----GRVDV-----SAHAKGNLHAIRK- 227  
 QY 546 SAVLTNAPYLLNDCDHYINNSKAIKEAMCFMADPLGKVCYVQFQRFQDIDRHRYA 605  
 Db 228 --RTSGDYIILDCD-I-PRAFLQIAMGM--VADRKIALQTPHHFYSFD----- 274  
 QY 606 NRVNVPFINMKGLDGIQPIYVGTGCVFRQALGYDAPRTKRPSTCNCWPRKCC 665  
 Db 275 -----PFQNNLAVGY-----RTPPE----- 289  
 QY 666 CCFGRKOKTTKPTKEKKLLFPKKEENQPAYALGEIDBAQEAENEKAGIVNQKLE 725  
 Db 290 ---GN-----LFY-----GVIOD---G 300  
 QY 726 KKFQGSVFVTSTLLENGTLKSPASPLKEAIVHISGVEDXTDMGKEIGMTYGSYTE 785  
 Db 301 NDFWDATFF-----CCSCAILREAIESIG-----GPAVEITYE 334  
 QY 786 DILGFKMHCWMSIYC-IPKRYVAFKGSAPLNSDRLOVLRWALGSIETFFSNHCPLM 844  
 Db 335 DAHTALRMQRGWSVAYLRIP---VASGLATERLTTHIGQMRMARAGMIQIFRYDNPML- 390  
 QY 845 YGVGGGLKFLRRFSYINSIYVPMTSIPLAYCTLPALCLGKFTPELNVASIPMSL 904  
 Db 391 ---GRGLGGRCLTSLMSTSPFAIPRVITLASPLATLPFGQNIIAAPLAVAYALPH 447  
 QY 905 FICIFATSIEMRWSGVGDIDMWRNDEFW-VIGVSSHLPFVFGQLKVIAGVDTSTFTV 963  
 Db 448 MFSIATA-----AKVNGMR-YSPFWSVEYETMALFLVRYITVILLPPSKGKFNV 498  
 QY 964 SKGG--DDEESELATFK--WTLILLPPTLLINPFGVAVGSNAINNGYESKGPLFG 1018  
 Db 499 EKGAVLEEEEDLGAIPNIIFAIIMGGILIGFELIVRNQDVIARNAY----- 550  
 QY 1019 KLFFAFVTVHLYPELKLGVGRNR- 1043  
 Db 551 -LINCAMALISLIIIFAAIVGRETK 575

RESULT 9  
 BCSA\_XANAC  
 ID\_BCSA\_XANAC STANDARD; PRT; 729 AA.

AC P58932: (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
 GN BCSA OR XAC3518.  
 OS Xanthomonas axonopodae (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Camavan F., Cardozo J., Chambergo F., Clapina L.P.,  
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RA "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities";  
 RL Nature 417:459-463(2002).  
 CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
 CC uridine 5', diphosphate glucose to cellulose, which is produced as  
 CC an extracellular component for mechanical and chemical protection  
 CC (by similarity).  
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP  
 CC + ((1,4)-beta-D-glucosyl)(N+1).  
 CC -1- COFACTOR: Magnesium (by similarity).  
 CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
 CC (c-di-GMP) (by similarity).  
 CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- DOMAIN: There are two conserved domains in the globular part of  
 CC the protein: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXXRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processive glycan chain  
 CC of the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: AEO12000; AAM38361.1; -  
 CC DR InterPro: IPR001173; Glyco trans 2.  
 CC DR Pfam: PF00535; Glycos transf 2; 1.  
 CC KM Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 CC Transmembrane; Inner membrane; Complete proteome;  
 CC FT DOMAIN 151 244 CATALYTIC SUBDOMAIN A.  
 CC FT DOMAIN 321 381 CATALYTIC SUBDOMAIN B.  
 CC FT TRANSMEM 30 50 POTENTIAL.  
 CC FT TRANSMEM 110 130 POTENTIAL.  
 CC FT TRANSMEM 171 191 POTENTIAL.  
 CC FT TRANSMEM 405 425 POTENTIAL.

FT TRANSMEM 427 447 POTENTIAL.  
 FT TRANSMEM 520 540 POTENTIAL.  
 FT TRANSMEM 549 569 POTENTIAL.  
 FT TRANSMEM 610 630 POTENTIAL.  
 FT ACT SITE 193 193 POTENTIAL.  
 FT ACT SITE 337 337 POTENTIAL.  
 FT BINDING 240 240 SUBSTRATE (POTENTIAL).  
 FT BINDING 242 242 SUBSTRATE (POTENTIAL).  
 SQ SEQUENCE 729 AA, 80915 MW, B9C0BB995E795B1 CRC64;  
 Query Match 3.9%; Score 230.5; DB 1; Length 729;  
 Best Local Similarity 18.9%; Pred. No. 2.5e-09;  
 Matches 139; Conservative 82; Mismatches 185; Indels 33; Gaps 26;  
 QY 299 PALMTLSVICIEMFAMSWILDOFPKMPRIERTYLDRLSLRDEKQSPQLAPIDFVST 358  
 DB 110 FLTGGLGALAEYALVAVLVIGFYGLMPLNRKPVPLPADQLMPSVDFVIFPT 161  
 QY 359 VDLKEPPLVNTNTLSLSVDYPVDXKSCVYSDGAMLTREALSETSEPAKKVVPFK 418  
 DB 162 VN--EPISVRTVLAASVIDWPAKITTILDDG----- 194  
 QY 419 RYNIERPAPEWYFOQKIDYLDKQVANFVREBRAMKREYBEKVRINALVAKQVPEBG 478  
 DB 195-----RR-----DEPR----- 200  
 QY 479 WTMQDGTWPWGNVDRHDMQIQLFGQSGGLDCENGLPRLVYVSREKRPQYNNHKKAGA 538  
 DB 201-----AFCAEVG-----INVTRTN-----MAAKXGN 223  
 QY 539 NMAALRVASVLTNARYLNDLDCDHYNNSKAIKEM-CPMNDPLGKVCYQVQRF-- 595  
 DB 224 INAAK----KCSGDYVAIFDGDH-IPTRSFLQVAMGFHLD---TTLAVQMHPYFES 274  
 QY 596-----DGIDRHDRYANRVNVFPDINMKGLDGIQGIYVGTGCVFRRQALYGVDAKTKP 650  
 DB 275 PDPFERNLDTHKVNBEGELFTGLIQDGDNDQMAITFGGSCVIRTLL----- 323  
 QY 651 PSRTNCWPKWCFCCCFGNRKQKTKTKTEKKLLFPKKEBNSPAYALGEIDEAAG 710  
 DB 324----- 323  
 QY 711 AENRAGIYNQKLEKKGQSSVFTSLNGLTKASPASLKEAIIHISGYEDKT 770  
 DB 324----- 323  
 QY 771 DWKEIGMT-YGVTEEDILITGFQKHCHGRSICYIPKRVAFGSAPLNLSPRLHQVLRMA 829  
 DB 324---EEVGVAAVETVEDAHTALKRGRTAYLAVPQA--GLATESLGHVAORIRMA 378  
 QY 830 IGSIEIFPSNHCPLWYGVGGGLKFLERPSYINSIVYPTWTSIPLAVCTLPALICLITGKFI 889  
 DB 379 RMAQIARIDN-PL--LGRGLKLSQRCLYANLAMPFYGVPRITLYLAPLAVLFGFAHV 434  
 QY 890 TRELNNVASLWMSLFLICFA-TSLLEMRNSGVGIDMMR---NEQF---WIGGVS 939  
 DB 435-----IQASALMILAYALPHILOANLTNLRVQSPFRHLNMEVETTLAWYI----- 481  
 QY 940 SHLFAVFOGLKLVINGVDT-SPTVTSKGG-----DDEPSELYEFKMTTLLIPTTLL 992  
 DB 482-----FRPTVALNPKLKGKENVTPKGLVARSYFDQAIAKPYLF-----LTL 524  
 QY 993 INFIGVAVGSNAING 1009  
 DB 525 LNVGVAVGVLRLLIYV 541  
 RESULT 10  
 ID BCAA ACXY STANDARD; PRT; 1518 AA.  
 AC Q9RB02;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative cellulose synthase 2 [includes: Cellulose synthase  
 DE catalytic subunit (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding  
 DE domain (Cellulose synthase 2 regulatory subunit)].  
 GN BCSAB11-A.  
 OS Acetobacter xylinus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
 OC Acetobacteraceae; Gluconacetobacter.  
 OX NCBI\_TaxID=28448;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 7664 / IFO 13693;  
 RX MEDLINE=9310341; PubMed=10382968;  
 RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,  
 RT Inoue Y.;  
 RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM  
 RT 7664: implication of a novel set of cellulose synthase genes.";  
 RL DNA Res. 6:109-115(1999).  
 CC -1 CARBOLYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP  
 CC + ((1,4)-beta-D-glucosyl)(N+1).  
 CC -1 PATHWAY: Bacterial cellulose biosynthesis.  
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1 DOMAIN: There are two conserved domains in the globular part of  
 CC the catalytic subunit: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXKRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -1 SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE  
 CC -1 GLYCOSYLTRANSFERASE FAMILY 2.  
 CC -1 SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ACSE/BCSB  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AB015803; BAA77593.1; -;  
 DR InterPro: IPR003919; CellSynth\_A.  
 DR InterPro: IPR003920; CellSynth\_B.  
 DR InterPro: IPR001173; Glyco\_trans\_2.  
 DR Pfam: PF03170; BcsAB; 1.  
 DR Pfam: PF00535; Glycos\_transf\_2; 1.  
 DR PRINTS: PR01439; CELSINTHASE.  
 DR PRINTS: PR01440; CELSINTHASE.  
 KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 KW Transmembrane; Inner membrane.  
 FT DOMAIN 1 731  
 FT DOMAIN 732 1518  
 FT DOMAIN 144 237  
 FT DOMAIN 314 374  
 FT TRANSMEM 24 44  
 FT TRANSMEM 71 91  
 FT TRANSMEM 105 125  
 FT TRANSMEM 105 125  
 FT TRANSMEM 404 424  
 FT TRANSMEM 427 447  
 FT TRANSMEM 465 485  
 FT TRANSMEM 514 534  
 FT TRANSMEM 543 563  
 FT TRANSMEM 1481 1501  
 FT TRANSMEM 186 196  
 FT ACT\_SITE 330 330  
 FT BINDING 233 233  
 FT BINDING 235 235  
 SQ SEQUENCE 1518 AA; 166404 MW; 0BC99B35B6DB4543 CRC64;

Query Match 3.7%; Score 218; DB 1; Length 1518;  
 Best Local Similarity 18.0%; Pred. No. 6.4e-08;  
 Matches 145; Conservative 88; Mismatches 223; Indels 340; Gaps 25;  
 QY 275 IIRLVVLPFFPHRYVHVPNDAPALML-----ISVICIWRAMSHILDPPKRPPIE 328  
 DB LMMSLVLSLVYIVRLTATVQ--FSNVLQPALAVLLILAEAYALMTLCLSYFQMAWPLR 132  
 QY 329 RETYIDRLSLRFDKQGPSQALPIDFVSTVDPLKEPPLVYTNVTLSTLSVDYVDKXSC 388  
 DB RREH---PLPEDAQWPS---VDVFVPSYN---BELSLVRSYVLGALDIDWPAIDLNV 181  
 QY 389 VYSDGAAMLTFEALSETSEFAKKVVPCKRYNIEPRAPEWYFOQKIDYLDKQKVAANVR 448  
 DB YILDDG----- 187  
 QY 449 ERRAKREYEPKRYINALVAKQVPEEGTMDGTGTPGNVNRDHPGMIQVFLGOSGG 508  
 DB 188 ---RRKAFHFAVAGA----- 201  
 QY 509 LDCEGNELPRLVYYSREKPGVNHKKAGANALVRSVAULTNAPYLLNDCDHYNNK 568  
 DB 202 ---GYIIRAE---NNHAKAGNHLAV---TDSPPAVIFPCDHPVPTRGF 242  
 QY 569 AIKEAMCFMMDPLGKVCYVQPFQDFGIDRHDRYANRVVFPDINKGLDGIGPIYV 628  
 DB 243 LRRTIGMMADP---NLALQTPHHFYAPRPPQ---NLAG--GMHPV--- 282  
 QY 629 GTGCVFRQALYGDAPRTKPPSRTCNCWPKWCFCGCCFQNRKOKTKTKTEKKLLF 688  
 DB 283 PEGNMFYGLVDGND-----FMDATFCGSC----- 308  
 QY 689 FKKEENQPAALGSDIDAPGAENKAGIYNQCKLEKFGQSSVFTSTLLENGTILKS 748  
 DB 309 ----- 308  
 QY 749 ASPALKEALHVISCGYEDKTDMGKEIGWTVGSYTEDILGPKHCHGMRSYICPRKV 808  
 DB 309 ---AIRREAVWGIC-----GFATRYTEALHAKMQRGKGTAYIRBPLA 352  
 QY 809 AFKSGAPLNLDRLOHVLQVLMALGSIIEFNSHCPLMYGGGLKFLRPSYINSIVYPT 868  
 DB 353 A--GLATRLILHIGQVRMARGMQIWRLDNPM---GAGLRBQGLCTLSASHFLF 406  
 QY 869 SIPLAYCTLPALICLTGKFTTPELANVASLWMF-SLFCIFANISILERRSGVIGIDW 927  
 DB 407 AIPRLTFVSPPLAFLECONIIAASPLAISVALPHISVYITLSRIEGRW----- 457  
 QY 928 RNEQFM-VIIGVSHLFAVFOGLKVIAGVDTSPFVTSKGGDEFSFLYFKTTLILP 986  
 DB 458 -RYSFWSRIYETSLALFLVRIITVTLDPHKGKFNVDKGG----- 497  
 QY 987 PTTLLNFIQVAVGASNAINGESWGPFLGKLFAPFVIVLHLPFKGLVGRQNRTP 1046  
 DB 498 -----LARGYFDMAYPNVILAQVLAAL---LRGFG----- 529  
 QY 1047 IYIVWSI-----LASIFSILWRI 1066  
 DB 530 --IYVQFHDRLALOSFIINTLWVI 552  
 RESULT 11  
 BCS\_ACEXY STANDARD; PRT; 1518 AA.  
 AC 09WK75;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative cellulose synthase 3 [includes: Cellulose synthase  
 DE catalytic subunit (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding  
 DE domain (Cellulose synthase 3 regulatory subunit)].  
 GN BCSAB11-B.  
 OS Acetobacter xylinus.





```

Db 151 VDFEPTYN---BELSLVRLTVLGLSLGIDMPREKRVHIIIDG----- 190
Qy 412 KWPFPCKRYNIEPRADPEWYFOQKIDYLDKQVAAFNREBRAMKREYEEFKVRIINALVAKA 471
Db 191 -----RRPEF-----AAFAAE----- 201
Qy 472 QKVPBEGWMTQDTPWPGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSEKRPQYN 531
Db 202 -----CGAN-----YIARPT-----N 212
Qy 532 HHKAGAMALVRSVAULTAPRLNLDCHYINNSKAIKEMCGFMV-DLLGKVCYGV 590
Db 213 EHKAGALNTAIG---HTDGDYILTPDCD-VPTRAFLOLTGMWVEDP---KIALMQ 263
Qy 591 FPOFPGDIDRDRYANRNVFFDINMKGLDGIQGIYVGTGCVFRRQALGYDAPKTKP 650
Db 264 TPNHFTSPD-----PFGRLSLAGY-----RTP 285
Qy 651 PSRTCNCPKWCFOCCCFGRKOKTKTKTEKKLLFPKKEENQSPAYALGEIDEAPG 710
Db 286 PE-----GN-----LFY----- 292
Qy 711 AENKAGIYNQKLEKKGSSVFTSLTENGSTLKSAPSLKEAIVHISQYEDKT 770
Db 293 -----GVQVD---GNDPMDATFF-----CGSCALTRTAIEQIG----- 323
Qy 771 DWGKEIGWYGVTEIDITLGFKNHCHGMSIYICPKFAVFGS-APLNLSDRLHGYLRMA 829
Db 324 -----GFATQYLTEDAHTLAKQRLGMSYAL--RIPLAGGLATERLLHIGQVRMA 374
Qy 830 LGSIEIFPSNHCPLWYGGGLKLEPFYSINSIVPWTISIPLLAYTTLPALCLLNGKFI 869
Db 375 RGMLOIFRIND-PL---FGRGLSGWGRCLCYLSANTSEFLFVAPVIFLPSSEIARLFFQONI 430
Qy 890 ---TPELNNAVSLWMSLFCIFATSLTLEMRWSVG---IDDMWRNEQFM-VYGGVSSH 941
Db 431 IAAFP-----LALLAYALPHV-FHAAVGAASKINQGMV-YSPWSEVYETMA 474
Qy 942 LFAVFOGLKVIAGVDTSTVTSKGDDE---EFSLEYTFKMTLLIPPTLLILNFI 997
Db 475 LFLVRVITVTLSSRSRKFVTDKGLLEKGYDPLGAVY-----PHIILGLIMFG 525
Qy 998 VVAGVSAINNNGE-SWGPL-----FGKLFPAFVIVHLYPFLKGL-VGRQ-----NR 1043
Db 526 LARGV-----YELSPGHLDQIAERAYLLNSAMAMSLIILIAIAVGRETOQKNSHR 578
Qy 1044 TPTITIV 1050
Db 579 IPATIPV 585

RESULT 13
BCA3_ACEXY STANDARD; PRT; 745 AA.
AC Q9W6L1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase 1 catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSAT.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=28448;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 7664 / IFO 13693;
RX MEDLINE=9310341; PubMed=10382968;
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA Inoue Y.,
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT 7664: implication of a novel set of cellulose synthase genes.";
RL DNA Res. 6:109-115(1999).

```

```

CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC uridine 5'-diphosphate glucose to cellulose. The thick cellulostic
CC mats generated by this enzyme probably provide a specialized
CC protective environment to the bacterium (By similarity).
CC -1- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl] (N) = UDP
CC + [(1,4)-beta-D-glucosyl] (N+1).
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC (c-di-GMP) (By similarity).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (periplasm).
CC -1- DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB015802; BAA77585.1; -.
DR InterPro: IPR003919; CellSynth_A.
DR InterPro: IPR001173; Glyco_trans_2.
DR Pfam: PF00535; Glycoe_transf_2; 1.
DR PRINTS: PR01439; CELSTNTHASR.
KM Cellulose biosynthesis; Transferase; Glycosyltransferase;
KM Transmembrane; Inner membrane.
FT DOMAIN 147 240 CATALYTIC SUBDOMAIN A.
FT TRANSMEM 317 377 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 430 450 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
FT TRANSMEM 515 535 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 649 669 POTENTIAL.
FT ACT_SITE 189 189 POTENTIAL.
FT ACT_SITE 333 333 POTENTIAL.
FT BINDING 236 236 SUBSTRATE (POTENTIAL).
FT BINDING 238 238 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 745 AA; 83518 MW; 57EA0457A226F815 CRC64;

Query Match 3.3%; Score 194; DB 1; Length 745;
Best Local Similarity 18.7%; Pred. No. 1.5e-06;
Matches 156; Conservative 88; Mismatches 221; Indels 370; Gaps 36;

Qy 276 IIRLVVLFPPHYRMRHVNDAPALMLISVCEIMFAMSWILDG---FPKVP----- 326
Db 53 MIYALVCVAVFPIVGRKRSRRQVFLVLSALVSLRY-LTWRLTRLDFTWTGILGVT 111
Qy 327 -IEREYV-LDRLSLFDEKGPOLAP-----IDPFVSTVDPLKEPPLVTNT 372
Db 112 LLLAEIYALYMLFLSYFQITISFLHRAPLPLPANPDEWPIVDIFITPD---EALSLVRLT 168
Qy 373 VLSILSVDPVDKVCYSDDQAMLTPEALSTSEFAKWWPFCRRNIEPRADPEWYFQ 432
Db 169 VIGALGIDMPDPKRVNYIIDDG----- 190
Qy 433 QKIDYLDKQVAAFNREBRAMKREYEEFKVRIINALVAKQKVPBEGWMTQDTPWPGNNV 492
Db 191 -----RR-----EEFARFAACGAR----- 205

```



QY 493 RDHPGMIQVFLGSGGGLDCEGNELPRLVYVSREKPRGVNHKKAGAMALVVSALVINA 552  
 Db 206 -----YIAPD-----NAHAKAGNLNAIK-----HTTG 229  
 QY 553 PYLLNLDCDHYNNSKAIKEAMCFMMDPLLGKVCVYVQPEQEDGIDRDRANRVVFF 612  
 Db 230 DHLLILDCDH-IPTRAFILQISMGMWVS--DSNIALLOTPHHFYSPD----- 272  
 QY 613 DINMGKLDGIQGIPIYVGTGCVRRQALYGDAPKTKKPSRTGNCWPKWCFCCCGNKK 672  
 Db 273 -----PQRNLNAVY-----RTPPE-----GN-- 289  
 QY 673 OKTKTKPKTEKKLLPFKKENQSPAYALGEIDAAPGAENKAGIVNOCKLEKKGQSS 732  
 Db 230 -----LFY-----GVICD-----GNDFWDT 305  
 QY 733 VEVSTLTLENGTILKASAPSLIKENIHVISCYEDKTDWGKEIGMIVSYTEDILTGK 792  
 Db 306 FF-----CGSCAILRRKRIEIG-----GFATETVEDAHYALR 339  
 QY 793 MECHGWRSIYC-IPKRVAFKSGAPLNLSDRLHQVLRMALGSIIEFSSNHCPIWYGGGL 851  
 Db 340 MQRKRWSTAYLRIP--LASGLATERLITHIGQRMRMAGMIOFRVNDPML-----GSGI 392  
 QY 852 KELEERSYINSIVYPTWSIPLLAYCTLP-----AICLLGKFTPELNNVASL 899  
 Db 399 KLGGRLCYLSAMTSFPFAPRVIPLASPLAFPFSGNIIAASPLANGVYAPHM----- 446  
 QY 900 WMSLFCIFATSIEMRWGSGVIDDWTNEQFWITGVGSHLPAVFOGLKVIAGVDT- 958  
 Db 447 -FHSI-----ATA-----AKVNGMR-YSPW-----SEYEVVVALFVRVYIVTM 485  
 QY 959 -----SFTVTSKGG--DDEPSELTYFPMKTTLLPPTLLLNPIFGVAVGYNMINNGY 1010  
 Db 486 LFPSGKGVNTEKGVLEKEEFDLTATY-----PNIIFAIMALGLRG----- 529  
 QY 1011 ESMGPIFGKLPFAFWIVHLYPFLKGLVGRONRTPTIVVWS-----ILLASIFSL 1061  
 Db 530 -----LYALIFQHLDI-----SERAYVLANCIWVSIVSLIIMAVISV 566

RESULT 14  
 CBL\_HUMAN  
 ID CBL\_HUMAN STANDARD; PRT; 906 AA.  
 AC P25681;  
 DT 01-AUG-1991 (Rel. 19, Last sequence created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE CBL E3 ubiquitin protein ligase (EC 6.3.2.-) (Signal transduction  
 protein CBL) (Proto-oncogene c-CBL).  
 GN CBL OR CBL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91232862; PubMed=2030914;  
 RA Blake T.J., Shapiro M., Morse H.C. III, Langdon W.Y.;  
 RT "The sequences of the human and mouse c-cbl proto-oncogenes show  
 RT v-cbl was generated by a large truncation encompassing a proline-rich  
 RT domain and a leucine zipper-like motif."  
 RL Oncogene 6:653-657(1991).  
 RN [2]  
 RP FUNCTION.  
 RX MEDLINE=9445925; PubMed=10514377;  
 RA Joazeiro C.A., Wing S.S., Huang H.-K., Levenson J.D., Hunter T.,  
 RA Liu Y.-C.;  
 RT "The tyrosine kinase negative regulator c-Cbl as a RING-type,  
 RT E2-dependent ubiquitin-protein ligase."  
 RL Science 286:309-312(1999).  
 RN [3]

RP INTERACTION WITH SLA AND ZAP70, AND MUTAGENESIS OF GLY-306.  
 RX MEDLINE=99380595; PubMed=10443770;  
 RA Tang J., Sawadikol S., Chang J.-H., Burakoff S.J.;  
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell  
 RT receptor signaling."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).  
 RN [4]  
 RP INTERACTION WITH SLA2.  
 RX MEDLINE=21553259; PubMed=11696592;  
 RA Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,  
 RA Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,  
 RA Yu S., Chan E., Wu X., Li C., Wolsetschlager M., Aversa G.,  
 RA Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,  
 RA Mancebo H.S.Y., Wu J.;  
 RT "Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel  
 RT inhibitor of antigen receptor signaling."  
 RL J. Exp. Med. 194:1263-1276(2001).  
 RN [5]  
 RP INTERACTION WITH CD2AP.  
 RX MEDLINE=21265017; PubMed=11067845;  
 RA Kirsch K.H., Georgescu M.M., Shishido T., Langdon W.Y., Birge R.B.,  
 RA Hanafusa H.;  
 RT "The adapter type protein CMS/CD2AP binds to the proto-oncogenic  
 RT protein c-Cbl through a tyrosine phosphorylation-regulated Src  
 RT homology 3 domain interaction."  
 RL J. Biol. Chem. 276:4957-4963(2001).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 47-350.  
 RX MEDLINE=99176421; PubMed=10078535;  
 RA Meng W., Sawadikol S., Burakoff S.J., Beck M.J.;  
 RT "Structure of the amino-terminal domain of Cbl complexed to its  
 RT binding site on ZAP-70 kinase."  
 RL Nature 398:84-90(1999).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 47-434 IN COMPLEX WITH ZAP70  
 AND UBE2L3.  
 RX MEDLINE=20419298; PubMed=10966114;  
 RA Zheng N., Wang P., Jeffrey P.D., Pavletich N.P.;  
 RT "Structure of a c-Cbl-UbcH7 complex: RING domain function in  
 RT ubiquitin-protein ligases."  
 RL Cell 102:533-539(2000).  
 CC -1- FUNCTION: Participates in signal transduction in hematopoietic  
 CC cells. Adapter protein that functions as a negative regulator of  
 CC many signaling pathways that start from receptors at the cell  
 CC surface. Acts as an E3 ubiquitin-protein ligase, which accepts  
 CC ubiquitin from specific E2 ubiquitin-conjugating enzymes, and then  
 CC transfers it to substrates promoting their degradation by the  
 CC proteasome. Recognizes activated receptor tyrosine kinases,  
 CC including PDGFR, EGF and CSF1, and terminates signaling.  
 CC -1- SUBUNIT: Associates with NCK via its SH3 domain. The  
 CC phosphorylated C-terminus interacts with CD2AP via its second SH3  
 CC domain. Binds to ZAP70 and UBE2L3. Interacts with adaptors SLA and  
 CC SLA2.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- DOMAIN: The RING-type zinc finger domain mediates binding to an E2  
 CC ubiquitin-conjugating enzyme.  
 CC -1- PTM: Phosphorylated on tyrosine residues.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 UBA domain.  
 CC -1- DATABASE: NAME=AcIas Genet. CytoGenet. Oncol. Haematol.;  
 CC WWW="http://www.infobio.gen.fr/services/chromancer/Genes/CBLID171.htm".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL, X57110; CAA40393.1; -  
 DR PIR; A43817; A43817.  
 DR PDB; 1B47; 27-APR-99.

DR PDB; 1FBV; 17-JUL-00.  
 DR PDB; 2CBL; 18-MAY-99.  
 DR Genew; HGNC:1541; CBL.  
 DR MIM: 165360; -.  
 DR GO; GO:0003700; F:transcription factor activity; TAS.  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR InterPro; IPR003153; Cbl\_N.  
 DR InterPro; IPR000449; Uba\_domain.  
 DR InterPro; IPR01841; Znf\_ring.  
 DR Pfam; PF02761; Cbl\_N2; 1.  
 DR Pfam; PF02762; Cbl\_N3; 1.  
 DR Pfam; PF02262; Cbl\_N; 1.  
 DR Pfam; PF06227; Uba; 1.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00165; UBA; 1.  
 DR PROSITE; PS00303; UBA; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS00089; ZF\_RING\_2; 1.  
 KW Ligase; Ubl conjugation pathway; Proto-oncogene; Zinc-finger;  
 KW Phosphorylation; 3D-structure.  
 FT ZN FING 381 420 RING-TYPE.  
 FT DOMAIN 357 476 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 477 688 PRO-RICH.  
 FT DOMAIN 648 906 INTERACTS WITH CD2AP.  
 FT DOMAIN 689 834 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 856 895 UBA.  
 FT MOD\_RES 700 700 PHOSPHORYLATION.  
 FT MOD\_RES 774 774 PHOSPHORYLATION.  
 FT MUTAGEN 306 306 G->E: ABOLISHES INTERACTION WITH ZAP70.  
 FT HELIX 53 70 BUT DOES NOT AFFECT INTERACTION WITH SLA.  
 FT TURN 71 71  
 FT HELIX 73 75  
 FT TURN 84 101  
 FT TURN 102 102  
 FT HELIX 106 110  
 FT TURN 111 111  
 FT HELIX 113 136  
 FT TURN 137 141  
 FT HELIX 143 144  
 FT TURN 146 168  
 FT HELIX 170 172  
 FT HELIX 176 178  
 FT HELIX 184 194  
 FT TURN 195 196  
 FT STRAND 199 201  
 FT TURN 202 212  
 FT HELIX 218 228  
 FT TURN 230 231  
 FT STRAND 235 237  
 FT HELIX 238 248  
 FT HELIX 251 253  
 FT HELIX 254 262  
 FT TURN 263 263  
 FT TURN 266 267  
 FT STRAND 268 271  
 FT HELIX 274 281  
 FT HELIX 282 284  
 FT TURN 285 286  
 FT TURN 288 289  
 FT STRAND 290 295  
 FT TURN 297 298  
 FT TURN 300 301  
 FT STRAND 303 308  
 FT TURN 310 310  
 FT STRAND 314 311  
 FT TURN 320 321  
 FT HELIX 324 333  
 FT TURN 334 335  
 FT STRAND 339 340  
 FT TURN 341 342  
 SQ SEQUENCE 906 AA; 99646 MW; 7D686B050204AD8F CRC64;

Query Match 2.4%; Score 139.5; DB 1; Length 906;  
 Best Local Similarity 19.4%; Pred. No. 0.025;  
 Matches 115; Conservative 62; Mismatches 194; Indels 221; Gaps 29;

QY 37 QVQCICGSDVGLAEGDFVACNECAFVNCDCY-EYRREGTQNCPCCKTRYKRLKCC 95  
 DB 379 QLCRIKCAEN-----DQDVKIEPCGHLMTCSLTJWSBSG-QGCFPCRC---EIKGTE 427  
 QY 96 -----RVTGDEEDVDLDNEFWMDHDSQSVASMLYGMYSYGRGDDPGARQA 146  
 DB 428 PIYVDFPDPFGSGSLRQGAEGASP-NYDDDDERADDTLFM--MKELAGAKYRPPSP 484  
 QY 147 FQLANPVELTNGQVDDIPEQHALVP---SFMGGGKRIHPPLPYADPSLPVQPRSD- 202  
 DB 485 FSWAPQASLPVPVPRDLDP--QRCVPSASALGTASKASGLHKQKPLPVPPTLBDL 542  
 QY 203 --PSKDLAAYGGSYAMKERNENMKQROERHQGNQGGGDDGDADLPLMDEARQQLSR 260  
 DB 543 PPPPPDPDPYVGA-----ESRRQRRPLPCTG-----DCPSRDKLP----- 579  
 QY 261 KIPLPSSQINPYRMIIRLVVLGPFPHYVWHVNDAPALMLISVCEIWFAMSMILDQ 320  
 DB 580 --PVPSSRLGD-----SWLPRP 594  
 QY 321 PPKMFPPIRETYLD-----RISLRFDEKQPSQLAP-----IDFVSTVDPBK 363  
 DB 595 IPK-VFVGAPSSDPWTGRELTNHSLPF---SLPSQMEPRPDVPRLGSTFSLDTSQMN 650  
 QY 364 BPPLV-----TTTVVLSLSVDYVNDKV---SCYSDQDAAM----- 397  
 DB 651 SPSLVGPECDBPKIKPSSANAIYSLARPLPVPPLPGEQCEGEDTEVYTPSSRPLRP 710  
 QY 398 -----LTFEALSETSEFAKWPFCFRYNIPEBPAPWYFOGKIDYLK 439  
 DB 711 LDTSSSRACDCDOQIDSCITYEAM-----YNISQASB--ITESTTGE 752  
 QY 440 DKVAANFV-----RERRANKREYEERKVRINALVAK-----AQKVEEGWMTQDCTPMPG 489  
 DB 753 GNLAAAHANTGPREESENEDDGVDVPPVPVAVLARLTSLDISNASSSGWMLSDGD--PT 810  
 QY 490 NNVRDHPGMIOVFLGSGGLDCGNELPRLVYVSEKRP---GYVHHKKAQA 538  
 DB 811 TNV-----TEGSQVP-----ERPPKPPRPRINSERKAGS 839

RESULT 15  
 GLNA\_SYN2 STANDARD; PRT; 473 AA.  
 AC P28605;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 RA Wagner S.J., Thomas S.P., Kaufman R.I., Nixon B.T., Stevens S.E. Jr.;  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).  
 GN GLNA.  
 OS *Synechococcus* sp. (strain PCC 7002) (*Agmenellum quadruplicatum*).  
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.  
 OX NCBI\_TaxId=32049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PR-6;  
 RX MEDLINE=93139025; PubMed=7678591;  
 RA Wagner S.J., Thomas S.P., Kaufman R.I., Nixon B.T., Stevens S.E. Jr.;  
 RT "The glna gene of the cyanobacterium *Agmenellum quadruplicatum* PR-6  
 is nonessential for ammonium assimilation.";  
 RL J. Bacteriol. 175:604-612(1993).  
 CC -L- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +  
 L-glutamine.  
 CC -L- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO  
 HEXAGONS.  
 CC -L- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -L- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z13965; CAA78366.1; -  
DR HSSP; P06201; ILGR.  
DR InterPro; IPR001691; GLN\_synth.  
DR InterPro; IPR004809; GLNA.  
DR InterPro; IPR001637; GLNA\_adenyltn.  
DR Pfam; PF00120; gln-synt\_1.  
DR Pfam; PF03851; gln-synt\_N; 1.  
DR ProDom; PD001057; Gln\_synt\_C; 1.  
DR TIGRFAMs; TIGR00653; GLNA; 1.  
DR PROSITE; PS00180; GLNA\_1; 1.  
DR PROSITE; PS00181; GLNA\_ATP; 1.  
KW Ligase.  
SQ SEQUENCE 473 AA, 53030 MW, 4956C7A1EDBC576 CRC64;

Query Match 2.0%; Score 116; DB 1; Length 473;

Best Local Similarity 22.5%; Pred. No. 0.58;  
Matches 75; Conservative 44; Mismatches 124; Indels 90; Gaps 19;

QY 373 VLSTLSVDYP-----VDKVSCT--VSDGAMLTREALSETS-----EPAKKW 413  
DB 19 IIDLFVLDLPGIWHQCSFYQDLDBASFYDGVFPDSSIRGWKAINESDMANVPDPTTAW 78  
QY 414 V-PCKRYNI-----EPRAPRW-----FOOKIDYLDKVKANFPRERAMKREYE 458  
DB 79 IDPFCKEKTLSLISIKKEPRTEWYSRDRSTIAOKAVDYL---APSGIGDTAYFGPEAE 134  
QY 459 EF---KVRINALVAKAKVPE--EG-WTWQDSTPWPGNVNRDHPGMIQVFLGQSGGLDCE 512  
DB 135 FFVFPDVAFPDQENKGFYYVDSEGRW--NSGRKEPGGLANKPGYKQGYFVPVPTDILQ 192  
QY 513 GNEFLRLVYVSREKRPGRYNNHKK--AGANNAL--VRVSAYLTNAPYLLNDCDHYI--NNS 567  
DB 193 DMRTMLTLMAKCGVPIEKHHHEVATGQNEIGFRFATLLKADYLMTY---KYVIKNVA 249  
QY 568 KAIKAMCFMMDLLGKVCYVQFPQRFIDRH-----DRYANRVVFPDI 614  
DB 250 RRYGRTVTFMPKPL-----FNDNGSGMHTHOSLMKEGQPLFWGDRYANLSQALH- 299  
QY 615 NMKGIDGIGPIYVGTGCVFRQALYGYDAKPT 647  
DB 300 -----YIG-GILKNAFALLAFSPST 319

Search completed: December 15, 2003, 13:03:23  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:00:29 ; Search time 45 Seconds

(without alignments)  
6227.670 Million cell updates/sec

Title: US-09-720-383C-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNRRLVVI.....DEFLAKDGPILLEEGCLDCN 1086

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP\_archaea:\*  
2: SP\_bacteria:\*  
3: SP\_fungi:\*  
4: SP\_human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP Vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_rvirus:\*  
16: SP\_bacterioplasmid:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5918	100.0	1086	10	Q9LLI3	Q9LLI3 zea mays (m
2	5408.5	91.4	1059	10	Q9LLI4	Q9LLI4 zea mays (m
3	5228	88.3	1094	10	Q9LLI2	Q9LLI2 zea mays (m
4	4333.5	73.2	1091	10	Q93XQ1	Q93XQ1 nicotiana a
5	4278.5	72.3	1084	10	Q48947	Q48947 arabidopsis
6	4262	72.0	1084	10	Q9FJ22	Q9FJ22 arabidopsis
7	4209.5	71.1	1088	10	Q9SJJ2	Q9SJJ2 arabidopsis
8	4204.5	71.0	1069	10	Q9FIB9	Q9FIB9 arabidopsis
9	4202	71.0	1081	10	Q65338	Q65338 arabidopsis
10	3984	67.3	1075	10	Q9LLI9	Q9LLI9 zea mays (m
11	3977.5	67.2	1074	10	Q9LLI8	Q9LLI8 zea mays (m
12	3967.5	67.0	1026	10	Q9SWM6	Q9SWM6 arabidopsis
13	3963.5	66.8	1032	10	Q9XHP6	Q9XHP6 arabidopsis
14	3955.5	66.7	1081	10	Q8GSM2	Q8GSM2 populus tre
15	3945.5	66.4	1067	10	Q48946	Q48946 arabidopsis
16	3929.5	66.4	1067	10	Q9XGX6	Q9XGX6 gossypium h

17	3911	66.1	1076	10	Q9LLI5	Q9LLI5 zea mays (m
18	3905.5	66.0	1077	10	Q9LLI6	Q9LLI6 zea mays (m
19	3898.5	65.9	1065	10	Q9FHK6	Q9FHK6 arabidopsis
20	3886	65.7	1042	10	Q8I649	Q8I649 populus x c
21	3880.5	65.6	1079	10	Q9LLI1	Q9LLI1 zea mays (m
22	3875.5	65.5	1065	10	Q48948	Q48948 arabidopsis
23	3835.5	64.8	1063	10	Q9AV71	Q9AV71 oryza sativ
24	3826	64.7	1129	10	Q8IKZ6	Q8IKZ6 mesocraetium
25	3810	64.4	1055	10	Q8GZM8	Q8GZM8 arabidopsis
26	3804.5	64.3	1043	10	Q9FNC3	Q9FNC3 arabidopsis
27	3782	63.9	1065	10	Q9SKC5	Q9SKC5 arabidopsis
28	3476.5	58.7	978	10	Q9AXK0	Q9AXK0 zinnia eleg
29	3469.5	58.6	974	10	P93155	P93155 gossypium h
30	3421	57.8	985	10	Q9C528	Q9C528 arabidopsis
31	3420	57.8	985	10	Q8LRF5	Q8LRF5 arabidopsis
32	3410.5	57.6	974	10	Q8W1W0	Q8W1W0 gossypium h
33	3332.5	56.3	978	10	Q8I368	Q8I368 populus tre
34	3329.5	56.3	958	10	Q9SN37	Q9SN37 arabidopsis
35	3318	56.1	821	10	Q9LLI7	Q9LLI7 zea mays (m
36	3165	53.5	939	10	Q943H3	Q943H3 oryza sativ
37	2807	49.1	685	10	P93156	P93156 gossypium h
38	2330.5	39.4	1181	10	Q9SRW9	Q9SRW9 arabidopsis
39	2326	39.3	1145	10	Q9M9M4	Q9M9M4 arabidopsis
40	2324.5	39.3	1104	10	Q8GJZ9	Q8GJZ9 populus tre
41	2321.5	39.2	1170	10	Q9LH27	Q9LH27 oryza sativ
42	2311	39.1	1145	10	Q9LFL0	Q9LFL0 arabidopsis
43	2293.5	38.8	1127	10	Q8W3F9	Q8W3F9 oryza sativ
44	2256	38.1	1111	10	Q9SZL9	Q9SZL9 arabidopsis
45	2253	38.1	1127	10	Q93XQ0	Q93XQ0 nicotiana a

## ALIGNMENTS

### RESULT 1

ID	Q9LLI3	PRELIMINARY;	PRT; 1086 AA.
AC	Q9LLI3	01-OCT-2000 (TREMBLrel. 15, Created)	
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel. 22, Last annotation update)		
DB	Cellulose synthase-7.		
GN	CESA-7.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	PACCAD clade; Panicoideae; Andropogoneae; Zea.		
OX	NCBI_TaxID=4577;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20398328; PubMed=10938350;		
RA	Holland N., Holland D., Helentjaris T., Dhugga K.S.,		
RA	Xococonote-Cazares B., Delmer D.P.;		
RT	"A comparative analysis of the plant cellulose synthase (Cesa) gene		
RT	family";		
RL	Plant Physiol. 123:1313-1324(2000).		
DR	EMBL; AF200531; AAF89967.1; -		
DR	InterPro; IPR005150; Cellulose_synth.		
DR	Pfam; PF03552; Cellulose_synth; 1.		
DR	SMART; SM00184; RING_1; RING_2; 1.		
DR	PROSITE; PS00089; ZF_RING_2; 1.		
SQ	SEQUENCE 1086 AA; 122608 MW; DD03C73ABD13E2EB CRC64;		

Query Match 100.0%; Score 5918; DB 10; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 MEASAGLVAGSHNRRLVIRRDGPGPKPPREONGVQCIQGDVGLAPGDPFVACNE 60
Db	1 MEASAGLVAGSHNRRLVIRRDGPGPKPPREONGVQCIQGDVGLAPGDPFVACNE 60
Qy	61 CAPVRCRCYERERREGTQNCPCQCTRYKLKGCORVTDGDEBDGVDDLDNEFFWMDGDS 120

```

Db 61 CAFFVCDCEYERREGTONCPOCKTRYYKRLKCGQRTYGEEDDGVDDLNEENMGHDS 120
Qy 121 QSVASMLYGHMSYGRGDPNGAPQAFQALPNPVLITNGQVDDIPPEQHALVPSFMGG 180
Db 121 QSVASMLYGHMSYGRGDPNGAPQAFQALPNPVLITNGQVDDIPPEQHALVPSFMGG 180
Qy 181 GKRIHPLPVADPSLPVQPRSMDSKDLAAYGYGSVANKEREMNKORQERHQTGNDGG 240
Db 181 GKRIHPLPVADPSLPVQPRSMDSKDLAAYGYGSVANKEREMNKORQERHQTGNDGG 240
Qy 241 DDGDDADPLMDARQOLSKRIPLPSSQINPYMIIIRLVYTGFFHYHVMHPVNDAPA 300
Db 241 DDGDDADPLMDARQOLSKRIPLPSSQINPYMIIIRLVYTGFFHYHVMHPVNDAPA 300
Qy 301 LMLISVCEIWMFMSWILDOFPKMFIERETYLDRSLRDKGQSPQLPIPFYSTVD 360
Db 301 LMLISVCEIWMFMSWILDOFPKMFIERETYLDRSLRDKGQSPQLPIPFYSTVD 360
Qy 361 PLKEPPLVTTNTVLISLVDYVDVKVSCYVSDGAAMLTEALSETSEFAKKWVPCKRY 420
Db 361 PLKEPPLVTTNTVLISLVDYVDVKVSCYVSDGAAMLTEALSETSEFAKKWVPCKRY 420
Qy 421 NIEPRABEYFOQKIDYLKQVANAFTREBRAMREYEEKRINLVAAQVYBEWGWT 480
Db 421 NIEPRABEYFOQKIDYLKQVANAFTREBRAMREYEEKRINLVAAQVYBEWGWT 480
Qy 481 MODGTPWGNVNRHPGMIQVFLQSGGLCEGNEELRLVYVSREKPGYNHKKAGAN 540
Db 481 MODGTPWGNVNRHPGMIQVFLQSGGLCEGNEELRLVYVSREKPGYNHKKAGAN 540
Qy 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVOPORFDGIDR 600
Db 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVOPORFDGIDR 600
Qy 601 HDRIYANRVYFPIINAKGLDGIQPIYVGTGCVFRDQALYGYDAPKTKRPSRTCMWPK 660
Db 601 HDRIYANRVYFPIINAKGLDGIQPIYVGTGCVFRDQALYGYDAPKTKRPSRTCMWPK 660
Qy 661 WCFCCCFGNRKOKKTKPKTEKKKLLFFKKEENQSAVALGIDEAPGAENKAGIYN 720
Db 661 WCFCCCFGNRKOKKTKPKTEKKKLLFFKKEENQSAVALGIDEAPGAENKAGIYN 720
Qy 721 QOKLEKFGQSSVFTSTLLENGSTLKASAPASLKEAIVISCGYEDKTDWKEIGMIY 780
Db 721 QOKLEKFGQSSVFTSTLLENGSTLKASAPASLKEAIVISCGYEDKTDWKEIGMIY 780
Qy 781 GSATEDILTFKMGHCHGMSIYCI PKRVAFKGAPLNLSDRLHOVLFWALGSTEIPFSNH 840
Db 781 GSATEDILTFKMGHCHGMSIYCI PKRVAFKGAPLNLSDRLHOVLFWALGSTEIPFSNH 840
Qy 841 CPLMYGGGGLKPLERFSYINSTIYPMWTSIPLAAYCTLPALCULTGKFIPELNNVSLM 900
Db 841 CPLMYGGGGLKPLERFSYINSTIYPMWTSIPLAAYCTLPALCULTGKFIPELNNVSLM 900
Qy 901 FMSLFCIFATSIILEMMSGVIGIDMMRNEQFWIIGVSSHLPFAVFOGLKVLAVDTSF 960
Db 901 FMSLFCIFATSIILEMMSGVIGIDMMRNEQFWIIGVSSHLPFAVFOGLKVLAVDTSF 960
Qy 961 TVTSKGGDDDEFSSELYFPKWTLLIIPPTLLILNFIIVAGVSNAINNGESWGPLFGKL 1020
Db 961 TVTSKGGDDDEFSSELYFPKWTLLIIPPTLLILNFIIVAGVSNAINNGESWGPLFGKL 1020
Qy 1021 FFAFWVIVHLVYPLKGLVGRNRTPTIYVWSILLASIFSLMWRIDPFLAKODGPLLEE 1080
Db 1021 FFAFWVIVHLVYPLKGLVGRNRTPTIYVWSILLASIFSLMWRIDPFLAKODGPLLEE 1080
Qy 1081 CGLDGN 1086
Db 1081 CGLDGN 1086

```

```

Q9LL14
ID Q9LL14 PRELIMINARY; PRT; 1059 AA.
AC Q9LL14;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Cellulose synthase-6.
GN CESA-6.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACMAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_Taxid=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Kocconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
family";
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200530; AAF89966.1; -
DR InterPro: IPR005150; Cellulose_synth.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synth; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B25232249 CRC64;

```

```

Query Match 91.4%; Score 5408.5; DB 10; Length 1059;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

```

```

Qy 33 EONGVCOJICGDDVGLAPGDPFVACNECAFVCRDCEYERREGTONCPOCKTRYKRLK 92
Db 3 QNRGVCQJCGDDVGNRPGEFVACNECAFICDCEYERREGTONCPOCKTRYKRLK 62
Qy 93 GCGRVYTGDEEDGVDDLNEFNW-DGHSQSVASMLYGHMSYGRGDPNGAPQAFQALNP 151
Db 63 GCGRVYTGDEEDGVDDLNEFNWSDKHDSQVLAESMLYGHMSYGRGADLDGVQFPHPI 122
Qy 152 NNPBLITNGQVDDIPPEQHALVPSFMGGGKRIHPLPVADPSLPVQPRSMDSKDLAAYG 211
Db 123 NNPBLITNGQVDDIPPEQHALVPSFMGGGKRIHPLPVADPSLPVQPRSMDSKDLAAYG 182
Qy 212 YGSVANKEREMNKORQERHQTGNDGGDDDDADPLMDARQOLSKRIPLPSSQINP 271
Db 183 YGSVANKEREMNKORQERHQTGNDGGDDDDADPLMDARQOLSKRIPLPSSQINP 242
Qy 272 YRMIIIRLVYLGFFHYHVMHPVNDAPALMLISVCEIWMFMSWILDOFPKMFIERET 331
Db 243 YRMIIIRLVYLGFFHYHVMHPVNDAPALMLISVCEIWMFMSWILDOFPKMFIERET 302
Qy 332 YLDRSLRDKGQSPQLPIPFYSTVDPLKEPPLVTTNTVLISLVDYVDVKVSCYVS 391
Db 303 YLDRSLRDKGQSPQLPIPFYSTVDPLKEPPLVTTNTVLISLVDYVDVKVSCYVS 362
Qy 392 DQGAAMLTEALSETSEFAKKWVPCKRNIEPRABEYFOQKIDYLKQVANAFTREBR 451
Db 363 DQGAAMLTEALSETSEFAKKWVPCKRNIEPRABEYFOQKIDYLKQVANAFTREBR 422
Qy 452 AMREYEEKRVIRINALVAKAKVPEEGWMTMODGTPWGNVNRHPGMIQVFLQSGGLDC 511
Db 423 AMREYEEKRVIRINALVAKAKVPEEGWMTMODGTPWGNVNRHPGMIQVFLQSGGHV 482
Qy 512 EGNELPRLVYVSREKPGYNHKKAGANALVRSVAVLTNAPYLNLDCDHYINNSKAIK 571
Db 483 EGNELPRLVYVSREKPGYNHKKAGANALVRSVAVLTNAPYLNLDCDHYINNSKAIK 542
Qy 572 EAMCFMMDPLGKVCYVOPORFDGIDHRIYANRVYFPIINAKGLDGIQPIYVGTG 631
Db 543 EAMCFMMDPLGKVCYVOPORFDGIDHRIYANRVYFPIINAKGLDGIQPIYVGTG 602

```

QY 632 CFFRRQALGYDAPAKTKKPPSRCTNCMPKCFCCCFGNRKOK---KTTKPKTEKKLLF 688  
DB 603 CFFRRQALGYDAPAKTKKPPSRCTNCMPKCFCCCFGNRKOKTKKTKSKPKFEKIKL- 661  
QY 689 FKKEENQSPAYALGEIDEAPGAENKAGI VNOQKLEKKFGQSSVFVSTLLENGTLLKS 748  
DB 662 FKKEQQAQAYALGEIDEAPGAENKAGI VNOQKLEKKFGQSSVFVASTLLENGTLLKS 721  
QY 749 AAPASILKKAHVISGVEDKTDWKEIGMIVGSTVEDILTGKPMCHQMRSTYICIPKRV 808  
DB 722 AAPASILKKAHVISGVEDKTDWKEIGMIVGSTVEDILTGKPMCHQMRSTYICIPKRV 781  
QY 809 AFGKSAPLNLSRDLHQLVRLMALGSIIEFNSHCPILVYGGGLKFLERFSYINSIYVPT 868  
DB 782 AFGKSAPLNLSRDLHQLVRLMALGSIIEFNSHCPILVYGGGLKFLERFSYINSIYVPT 841  
QY 869 SIPLLAYCTLPALICLLTGKFTPELNNVASLWMSLFCIFATSILEMNSGVGIDDMWR 928  
DB 842 SIPLLAYCTLPALICLLTGKFTPELNNVASLWMSLFCIFATSILEMNSGVGIDDMWR 901  
QY 929 NQOPWIVGVSLSLFAVFGGLKVLVGVDTSTFTVTSKGGDDEFESELYTFKMTLLIPT 988  
DB 902 NQOPWIVGVSLSLFAVFGGLKVLVGVDTSTFTVTSKGGDDEFESELYTFKMTLLIPT 961  
QY 989 TLLNLFIVGVSNAINGVSGPLFGKLFFAFWVIVHLVPLKGLVGRONRTPTIV 1048  
DB 962 TLLNLFIVGVSNAINGVSGPLFGKLFFAFWVIVHLVPLKGLVGRONRTPTIV 1021  
QY 1049 IVMSILASISFSLWVRIIDPLAKDGPPLLEBEGJLDCN 1086  
DB 1022 IVMSILASISFSLWVRIIDPLAKDGPPLLEBEGJLDCN 1059

## RESULT 3

Q9LL12 PRELIMINARY; PRT; 1094 AA.  
AC Q9LL12;  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)  
DE Cellulose synthase-8.  
GN CESA-8.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20398328; PubMed=10938350;  
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,  
RA Xocoostle-Cazares B., Delmer D.P.;  
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene family";  
RL Plant Physiol. 123:1313-1324(2000).  
DR EMBL; AF200532; AAF89968.1; -  
DR InterPro; IPR005150; Cellulose\_synth.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF03552; Cellulose\_synth; 1.  
DR SMART; SM00184; RING\_1.  
DR PROSITE; P550089; ZF\_RING\_2; 1.  
SQ SEQUENCE 1094 AA; 122575 MW; DFBGCI8F49D23F5E CRC64;

Query Match 88.3%; Score 5228; DB 10; Length 1094;  
Best Local Similarity 88.7%; Pred. No. 0;  
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

QY 1 MEASAGLVAGSHNRNELVIRRD--GDPGPKPRBQNGQVCICGDDVGLAPGDPFVA 57  
DB 1 MEASAGLVAGSHNRNELVIRRDRESGAAGGAARAPAP-CQICGDEVGVGDPFVA 59  
QY 58 CNECAFVPCRDCEYRREGTQNCPOCKTRRYLKKCCQVTDDEEDSGVDDIDNEN-WD 116

DB 60 CNECAFVPCRDCEYRREGTQNCPOCKTRRYLKKCCQVTDDEEDSGVDDIDNEN-WD 119  
QY 117 G-----HDSQVASESLYGHMSYGRGDPNGAPQALQPLNPVPLTNGQWDDIPPEQNAL 172  
DB 120 GAHEDDPQYVAESMLRAQMSYGRGD-- -APGGSVPVENPPLTNGQWDDIPPEQNAL 176  
QY 173 VPSFW---GGGGRJHPLPYADPSLPVOPRSDPESKDLAAYGVGVANKERENKQOE 229  
DB 177 VPSVNSGGGGGGRJHPLPYADPSLPVOPRSDPESKDLAAYGVGVANKERENKQOE 236  
QY 230 RMHQGTNDGGD-DGDDADPLMDBARQOLSRLKPLPSSQINPYMIIIRLVVLGFFPH 288  
DB 227 RLQHRSSGGGWDGDDADPLMDBARQOLSRLKPLPSSQINPYMIIIRLVVLGFFPH 296  
QY 289 YRVHPVNDAPALMLISYICELWPMMSVILDOFPWFPIEBETYLDRSLRDKGQPSQ 348  
DB 297 YRVHPVNDAPALMLISYICELWPMMSVILDOFPWFPIEBETYLDRSLRDKGQPSQ 356  
QY 349 LAPIDFVSTVDPLEKPLVTTNTVLSTLSVDYPAVDKYSYVSDGAAMLTPEALSTSE 408  
DB 357 LAPIDFVSTVDPLEKPLVTTNTVLSTLSVDYPAVDKYSYVSDGAAMLTPEALSTSE 416  
QY 409 FAKKVPPECKRYNIEPRAPENYFOQKIDYLKDKVANFVREBRAMKREYEEKVINLV 468  
DB 417 FAKKVPPECKRYNIEPRAPENYFOQKIDYLKDKVANFVREBRAMKREYEEKVINLV 476  
QY 469 AKAQVPEBGMTQDGTWPGNNVDHPGMIQVFLVGGGGLDCBGNELPRLVYSREKRP 528  
DB 477 AKAQVPEBGMTQDGTWPGNNVDHPGMIQVFLVGGGGLDCBGNELPRLVYSREKRP 536  
QY 529 GYNHKKKAGANVALRVSAVLTNAPYLNTLDCDHYNNSKAIKEMCMTPMDPLKRYCY 588  
DB 537 GYNHKKKAGANVALRVSAVLTNAPYLNTLDCDHYNNSKAIKEMCMTPMDPLKRYCY 596  
QY 589 VQFPQRPDGDIDRHDYANRVVFEEDINNKGLDGIQPIYVGVGVCFRRQALGYDAPKTK 648  
DB 597 VQFPQRPDGDIDRHDYANRVVFEEDINNKGLDGIQPIYVGVGVCFRRQALGYDAPKTK 656  
QY 649 KPSPSTCNCMPKCFCCCFGNRKOKTKTKPKTEKKLLFPKKEENQSPAYALGEIDEA 708  
DB 657 KPSPSTCNCMPKCFCCCFGNRKOKTKTKPKTEKKLLFPKKEENQSPAYALGEIDEA 716  
QY 709 PGAENKAGIYNOQKLEKKFGQSSVFVSTLLENGTLLKSASPAALLKKAHVISGVED 768  
DB 717 PGADIEKAGIYNOQKLEKKFGQSSVFVASTLLENGTLLKSASPAALLKKAHVISGVED 776  
QY 769 KTDWKEIGMIVGSTVEDILTGKPMCHQMRSTYICIPKRVAFKGSAPLNLSDRLHQLVRL 828  
DB 777 KTDWKEIGMIVGSTVEDILTGKPMCHQMRSTYICIPKRVAFKGSAPLNLSDRLHQLVRL 836  
QY 829 ALGSIIEFFSNHCPILVYGGGLKFLERFSYINSIYVPTWSIPLAYCTLPALICLLTGK 888  
DB 837 ALGSIIEFFSNHCPILVYGGGLKFLERFSYINSIYVPTWSIPLAYCTLPALICLLTGK 896  
QY 889 ITPELNNVASLWMSLFCIFATSILEMNSGVGIDDMWRNQOPWIVGVSLSLFAVFG 948  
DB 897 ITPELNNVASLWMSLFCIFATSILEMNSGVGIDDMWRNQOPWIVGVSLSLFAVFG 956  
QY 949 LKAVAGVDTSTFTVTSKGGDDEFESELYTFKMTLLIPTTLLNLFIVGVSNAINGV 1008  
DB 957 LKAVAGVDTSTFTVTSKGGDDEFESELYTFKMTLLIPTTLLNLFIVGVSNAINGV 1016  
QY 1009 GYESWGPFLGKLFPAFWIVHLVPLKGLVGRONRTPTIVVMSILASISFSLWVRIID 1068  
DB 1017 GYESWGPFLGKLFPAFWIVHLVPLKGLVGRONRTPTIVVMSILASISFSLWVRIID 1076  
QY 1069 FLAKDGPPLLEBEGJLDCN 1086  
DB 1077 FLAKDGPPLLEBEGJLDCN 1094

RESULT 4  
Q93XQ1

ID 093X01 PRELIMINARY; PRT; 1091 AA.  
 AC 093X01;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2003 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Cellulose synthase catalytic subunit.  
 CS EMBL.  
 GN Nicotiana glauca (winged tobacco) (Persian tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_TaxID=4087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Petiole;  
 RX MEDLINE=21196092; PubMed=11299383;  
 RA Dohlin M.S., De Melis L., Newbigin B., Bacic A., Read S.M.;  
 RT "Pollen Tubes of Nicotiana glauca Express Two Genes from Different  
 beta-Glucan Synthase Families";  
 RL Plant Physiol. 125:2040-2052(2001).  
 DR EMBL; AF304374; AAK49454.1; -  
 DR InterPro: IPR005150; Cellulose\_synth.  
 DR InterPro: IPR001841; Zn\_finger.  
 DR Pfam: PF03552; Cellulose\_synth. 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 DR SEQUENCE 1091 AA; 122682 MW; 3FB8DF94B19D519B CRC64;

Query Match 73.2%; Score 4333.5; DB 10; Length 1091;  
 Best Local Similarity 71.9%; Pred. No. 0;  
 Matches 794; Conservative 123; Mismatches 153; Indels 35; Gaps 12;

QY 1 MEASAGLVAGSHNRNELVIRDDGPGKPPREQNGVCGICGDVGLAFGDPFVAACNE 60  
 DB 1 MDTKGLVAGSHNRNEFVIVNADVGKVTSEKLSGQICQICGDEIEVTVDSGPFACNE 60  
 QY 61 CAPPVGDCYCYEYRRREGTQNCPOCKTRRYKLGKCGVYTGEBEEDVDLDNENMGHD- 119  
 DB 61 CAPPVGDCYCYEYRRREGTQNCPOCKTRRYKLGKCGVYTGEBEEDVDLDNENMGHD- 120  
 QY 120 --SQAASMLYGHMSYGRGDPNGA-----POAFQJNPNVPLLTNGQAVDDIPE 168  
 DB 121 YSEEAFFSRLL-----GRGTNNASGLTTPSEVDPA--LNEIFILITIGGDDITISAD 172  
 QY 169 QHALL-VPSFNGGSGKRIHPLRYADPSLVQPRSMDSKDLAAYGYSVANKEMENMKOR 227  
 DB 173 KHALIIPRFV-GGKGVKHPVPSD-SMSLPRPMDEPKDLAVYGYGVANKEMEDMKKK 230  
 QY 228 QERMGOT---GNDGGGDDG---DDADLPLMDEARQQLSKITLPSSQINPYMIITIRL 280  
 DB 231 QNDKLVVKKHGGKGGGNDDELDDPLPMDGGRQPLSRKLPISSSRSLSPYRLILVRL 290  
 QY 281 VVLFPPFHYRVWMPVNDAPALMLISVCEIWPAMSWMLDOPPKMPPIBRETVDRLSLRP 340  
 DB 291 AVVGLFFHYRITPNDAYALMLISICETWFAVSWTFDOPKMPPIVRETVDRLSLRK 350  
 QY 341 DKEGOSQALPIDFVSTVDPLKEPRLVTNTVLSILSVDPVYDKVSCVSDGAAMLT 400  
 DB 351 EKGKSGSLAPIDIFSTVDPLKEPRLITANTVLSILAVDPYDKVSCVSDGAAMLT 410  
 QY 401 EALSTSEFAKKVNPFRCKINIPRAPEWFOOKIDYLGKVAANFPRERAKREYEER 460  
 DB 411 EALSTSEFAKKVNPFRCKINIPRAPEWFOOKIDYLGKVAANFPRERAKREYEER 470  
 QY 461 KVALNLVAKAKVPEEGMTMDGTPWGNVVDHPMIVVFGSGSGGLCEGNEPLRLV 520  
 DB 471 KARINGLVAKVPEEGMTMDGTPWGNVVDHPMIVVFGSGSGGLCEGNEPLRLV 530  
 QY 521 YVSRERKPGYVHKKAGANALVVSALVTNAPYLNLDCDHYINNSKAIKAMCFMMD 580  
 DB 531 YVSRERKPGYVHKKAGANALVVSALVTNAPYLNLDCDHYINNSKAIKAMCFMMD 590  
 QY 581 LLGKVCYVOPRQFSGIDRHDKYANRNVFPIDMKGLDGIQGPITYGTCVFRQALY 640

DB 591 TSGKKICVQFPQFQDIDRHDRYSNRNVFPIIMKGLDGIQGPITYGTCVFRQALY 650  
 QY 641 GYDAKTKKPPSPRTGNCMPKWCFCGCCGNRKOKKTKPKTEKKLLFFKKEENSPAY 700  
 DB 651 GYDAKTKKPPSPRTGNCMPKWCFCGCCGNRKOKKTKPKTEKKLLFFKKEENSPAY 706  
 QY 701 LAEIDEAPGAENKAGIVNQKLEKFKQSSVPTSTLLENGTLKASPAISLKEAII 760  
 DB 701 LENIEGEGEGIDSEKATLMPQIKLEKFKQSSVPTSTLLENGTLKASPAISLKEAII 766  
 QY 761 VISCCEYEDTKWCKEITGMYGSTEEDILTGPMHCHGRSTYCIKRAVAFKSAADINSD 820  
 DB 767 VISCCEYEDTKWCKEITGMYGSTEEDILTGPMHCHGRSTYCIKRAVAFKSAADINSD 826  
 QY 821 RLHQVLRWALGSIETFFSNHCPWLYGCGGLKFLERFSYINSIVPMTSIPPLACTLPA 880  
 DB 827 RLHQVLRWALGSIETFFSNHCPWLYGCGGLKFLERFSYINSIVPMTSIPPLACTLPA 886  
 QY 881 ICLTGKFTPLPNNVASIEMNSLFCIPATSIEMRWSGVGIDDMWRNEQFVIGVSS 940  
 DB 887 VCLLTGKFTPLPNNVASIEMNSLFCIPATSIEMRWSGVGIDDMWRNEQFVIGVSS 946  
 QY 941 HLFVAFQGLKVIAGVDSFTVTSKGDDESESLYTPKWTLLPPTLLILNFTGVA 1000  
 DB 947 HLFVAFQGLKVIAGVDSFTVTSKGDDESESLYTPKWTLLPPTLLILNFTGVA 1006  
 QY 1001 GVSNAIINGYESMGLFGLFPAFVIVHLYPFLKGLVGRORPTIYIVWSILASIFS 1060  
 DB 1007 GVSNAIINGYESMGLFGLFPAFVIVHLYPFLKGLVGRORPTIYIVWSILASIFS 1066  
 QY 1061 LLMVRIDFPLANDGDLBEGGLDC 1085  
 DB 1067 LLMVRIDFPLANDGDLBEGGLDC 1090

## RESULT 5

ID 048947 PRELIMINARY; PRT; 1084 AA.  
 AC 048947;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Cellulose synthase catalytic subunit.  
 GN H-A OR T22R8.250 OR AT4G39350.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroside II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98111412; PubMed=9445479;  
 RA Artoli T., Peng L., Betzner A.S., Burn J., Witte W., Herth W.,  
 RA Camilleri C., Hotte H., Plazinski J., Birch R., Cork A., Glover J.,  
 RA Redmond J., Williams R.E.;  
 RT "Molecular analysis of cellulose biosynthesis in Arabidopsis";  
 RL Science 279:717-720(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Zimmermann W., Grueneisen A., Wambutt R., Bancroft I.,  
 RA Mewes H.W., Mayer K.F.X., Schueller C.;  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Zimmermann W., Grueneisen A., Wambutt R., Kalicki J., Woldmann P.,  
 RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.



```

RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF0271173; AAC39335.1; -
DR EMBL; AL050351; CAB3650.1; -
DR EMBL; AL161595; CAB0598.1; -
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR01841; Zn_fing.
DR Pfam; PF03552; Cellulose_synth; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
SQ SEQUENCE 1084 AA; 122068 MW; 2F9B2D16BD734B0 CRC64;

Query Match 72.3%; Score 4278.5; DB 10; Length 1084;
Best Local Similarity 71.3%; Pred. No. 0;
Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;

QY 1 MEASAGIVAGSHNRELVIIRDDGPGRPREONGVQICGDDVGLAPGDDPFAACNE 60
DB 1 MNTGRLTAGSHNNEFLVINADESARIRSVQELSGTCQICGDIELTVSELVACNE 60
QY 61 CAFVPCRCDCYERREBGTQNCPOCKTRYKRLKGCQRTVG-DDEBDGVDDLNEFMWDGD 119
DB 61 CAFVPCRCDCYERREBGTQNCPOCKTRYKRLKGSFRVGDDEBEDIIDLETFD-HGMD 119
QY 120 SOSVAESMLYGHMSYRGDGPNGAPQAFQALNNVBLITNGQWDDIIPPEQHALVPSFMGC 179
DB 120 PEHAEAALSRILNTRGGLDAP-----SGQIFLTVCEBDADMSYDRHALIVPSTG 174
QY 180 GGRKIHPIPYADPSLPVQPRSMDESKDLAAYGYGVANKEMENMKQKQ-BRM---HOT 234
DB 175 YGNRYPAFPDSSAPQAFARSPVKDIAEYGYGVANKEMENMKQKQ-BRM---HOT 234
QY 225 GNDGGG-----DDGDADPLMDEARQOLSRKIPILSSQINPRMTIIRLVLAGFFHYR 290
DB 225 GNNRGSGDDDELDDPMDDEGRQPSRKLPILSSINPRMTIIRLVLAGFFHYR 294
QY 291 VMHPVNDAFALMLISVCEIFPAMSWIIDQPKWPIRETYLDRLSLRFDKEGQPSOLA 350
DB 291 ILHPVNDAYGLMLISVCEIFPAMSWIIDQPKWPIRETYLDRLSLRFEKEGQPSOLA 354
QY 351 PIDFVSTVVDLKEPPLVTTNTVLSISVDYPRVDKVCYSDGGAAMITFEALSTSEPA 410
DB 355 PVDVFSVTDPLKEPPLITANTVLSISLAVDVPVDKVCYSDGGAAMITFEALSTSEPA 414
QY 411 KKMVPFCRNYIEPPAPBEMVFOQIKDYLDKVAAPFVPERAMKEVEEFPKRIYALAK 470
DB 415 KKMVPFCRNYIEPPAPBEMVFOQIKDYLDKVAAPFVPERAMKEVEEFPKRIYALAK 474
QY 471 AOKVPEEGMTQDGTMPGNNVRDHPMIQVFLGQSGLDCEGNEIPLLVYVSREKRPDY 530
DB 475 AOKVPEEGMTQDGTMPGNNVRDHPMIQVFLGSGVBDTDGNEIPLLVYVSREKRPDY 534
QY 531 NHHKKAAGAMALVRSVAVLTNAPLYLNTDCHYIINNSAIKEAMCFMMDPLIGKKVCYQ 590
DB 535 DHKKAAGAMNSLIRVSAVLTNAPLYLNTDCHYIINNSAIRESMCFMMDPOSGKKVCYQ 594
QY 591 FPOPRDGIIDRRYANRVVFPDINMKGLDGIQGIYVGTGCVPRQALYGVDAKTKRP 650
DB 595 FPOPRDGIIDRRYANRVVFPDINMKGLDGIQGIYVGTGCVPRQALYGVDAKTKRP 654
QY 651 PSRTCNCPKWCFCGCCFGNKKOKTKTRKTEKKLLPFKKKEENSRYVALGEIDEA--A 708
DB 655 PSRTCNCPKWCFCGCCFGNKKOKTKTRKTEKKLLPFKKKEENSRYVALGEIDEA--A 708
QY 709 FGAENKAGIVNQQLTEKKFGQSSVFTSTLENGSTLKSASPASLLEKALHVISGVED 768
DB 707 FVSNVKEKSEATQLTEKKFGQSSVFTSTLENGSTLKSASPASLLEKALHVISGVED 766
QY 769 KTDWKEIGWYIGSTEDILTGFKKHCHGMSIYICIPKRVAFKSGAPLNTSDRLHQLVIR 828
DB 767 KTDWKEIGWYIGSTEDILTGFKKHCHGMSIYICIPKRVAFKSGAPLNTSDRLHQLVIR 826

```

```

QY 829 ALGSEIFFSNHCMLWYGGGLKFLERFSYINSIVYPMISIPILAYCTLPALCLTKCF 888
DB 827 ALGSEVIFLSNHCIMWYGGGLKFLERFSYINSIVYPMISIPILAYCTLPALCLTKCF 886
QY 889 ITPELANNVASLPMFSLFCIPATSLILEMRMGVQIDDMWRNRCQVWIGVSSHFAVROG 948
DB 887 ITPELANNVASLPMFSLFCIPATSLILEMRMGVQIDDMWRNRCQVWIGVSSHFAVROG 946
QY 949 LLKVIAGVDSFTYTSKGGDDEEFSBELYTFKMTLLIPPTLLLNPIGVAGVSNAINN 1008
DB 947 LLKVIAGVDSFTYTSKGGDDEEFSBELYTFKMTLLIPPTLLLNPIGVAGVSNAINN 1006
QY 1009 GYBSWGLPFGCLFAFVWYVHLVPLKGLVGRQNRPTIYVMSILASISLWVRIDP 1068
DB 1007 GYBSWGLPFGCLFAFVWYVHLVPLKGLVGRQNRPTIYVMSILASISLWVRIDP 1066
QY 1069 FLAKDDPFLREGLDC 1085
DB 1067 FLAKDDPFLREGLDC 1082

RESULT 6
ID Q9FGF9 PRELIMINARY; PRT; 1084 AA.
AC Q9FGF9;
DT 01-MAR-2001 (TREMUREL.16, Created)
DT 01-MAR-2001 (TREMUREL.16, Last sequence update)
DT 01-OCT-2002 (TREMUREL.22, Last annotation update)
DE Cellulose synthase catalytic subunit.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Kaneo T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025637; BAB10307.1; -
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Zn_fing.
DR Pfam; PF03552; Cellulose_synth; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
SQ SEQUENCE 1084 AA; 122501 MW; 1520439A5053608C CRC64;

Query Match 72.0%; Score 4262; DB 10; Length 1084;
Best Local Similarity 71.1%; Pred. No. 0;
Matches 778; Conservative 126; Mismatches 170; Indels 20; Gaps 10;

QY 1 MEASAGIVAGSHNRELVIIRDDGPGRPREONGVQICGDDVGLAPGDDPFAACNE 60
DB 1 MNTGRLTAGSHNNEFLVINADESARIRSVQELSGTCQICRDEIELTVDSGFVACNE 60
QY 61 CAFVPCRCDCYERREBGTQNCPOCKTRYKRLKGCQRTVG-DDEBDGVDDLNEFMWDGD 119
DB 61 CAFVPCRCDCYERREBGTQNCPOCKTRYKRLKGSFRVGDDEBEDIIDLETFD-HGMD 119
QY 120 SOSVAESMLYGHMSYRGDGPNGAPQAFQALNNVBLITNGQWDDIIPPEQHALVPSFMGC 179
DB 120 PEHAEAALSRILNTRGGLDAP-----SGQIFLTVCEBDADMSYDRHALIVPSTG 174
QY 180 GGRKIHPIPYADPSLPVQPRSMDESKDLAAYGYGVANKEMENMKQKQ-BRM---HOT 234
DB 175 YGNRYPAFPDSSAPQAFARSPVKDIAEYGYGVANKEMENMKQKQ-BRM---HOT 234
QY 225 GNDGGG-----DDGDADPLMDEARQOLSRKIPILSSQINPRMTIIRLVLAGFFHYR 290
DB 225 GNNRGSGDDDELDDPMDDEGRQPSRKLPILSSINPRMTIIRLVLAGFFHYR 294
QY 291 VMHPVNDAFALMLISVCEIFPAMSWIIDQPKWPIRETYLDRLSLRFDKEGQPSOLA 350
DB 291 ILHPVNDAYGLMLISVCEIFPAMSWIIDQPKWPIRETYLDRLSLRFEKEGQPSOLA 354
QY 351 PIDFVSTVVDLKEPPLVTTNTVLSISVDYPRVDKVCYSDGGAAMITFEALSTSEPA 410
DB 355 PVDVFSVTDPLKEPPLITANTVLSISLAVDVPVDKVCYSDGGAAMITFEALSTSEPA 414
QY 411 KKMVPFCRNYIEPPAPBEMVFOQIKDYLDKVAAPFVPERAMKEVEEFPKRIYALAK 470
DB 415 KKMVPFCRNYIEPPAPBEMVFOQIKDYLDKVAAPFVPERAMKEVEEFPKRIYALAK 474
QY 471 AOKVPEEGMTQDGTMPGNNVRDHPMIQVFLGQSGLDCEGNEIPLLVYVSREKRPDY 530
DB 475 AOKVPEEGMTQDGTMPGNNVRDHPMIQVFLGSGVBDTDGNEIPLLVYVSREKRPDY 534
QY 531 NHHKKAAGAMALVRSVAVLTNAPLYLNTDCHYIINNSAIKEAMCFMMDPLIGKKVCYQ 590
DB 535 DHKKAAGAMNSLIRVSAVLTNAPLYLNTDCHYIINNSAIRESMCFMMDPOSGKKVCYQ 594
QY 591 FPOPRDGIIDRRYANRVVFPDINMKGLDGIQGIYVGTGCVPRQALYGVDAKTKRP 650
DB 595 FPOPRDGIIDRRYANRVVFPDINMKGLDGIQGIYVGTGCVPRQALYGVDAKTKRP 654
QY 651 PSRTCNCPKWCFCGCCFGNKKOKTKTRKTEKKLLPFKKKEENSRYVALGEIDEA--A 708
DB 655 PSRTCNCPKWCFCGCCFGNKKOKTKTRKTEKKLLPFKKKEENSRYVALGEIDEA--A 708
QY 709 FGAENKAGIVNQQLTEKKFGQSSVFTSTLENGSTLKSASPASLLEKALHVISGVED 768
DB 707 FVSNVKEKSEATQLTEKKFGQSSVFTSTLENGSTLKSASPASLLEKALHVISGVED 766
QY 769 KTDWKEIGWYIGSTEDILTGFKKHCHGMSIYICIPKRVAFKSGAPLNTSDRLHQLVIR 828
DB 767 KTDWKEIGWYIGSTEDILTGFKKHCHGMSIYICIPKRVAFKSGAPLNTSDRLHQLVIR 826

```

```

QY 294 PYNDAFALMISVIGCEWFMAMWILDOFPKMPRIERTYLDLSLFDKSGQPSQALPD 353
DB 297 PVKDAVALMISVIGCEWFMAMWILDOFPKMPRIERTYLDLSLFDKSGQPSQALPD 356
QY 354 FVSTYDPLKEPPLVNTNTVLSLSVDYPRVDKSCVSDGAMLTFFALSTSESPAKK 413
DB 357 VEVSTYDPLKEPPLVNTNTVLSLSVDYPRVDKSCVSDGAMLTFFALSTSESPAKK 416
QY 414 VPECKRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVAKOK 473
DB 417 VPECKRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVAKOK 476
QY 474 VPECKRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVAKOK 533
DB 477 VPECKRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVAKOK 536
QY 534 KXAGANALVAVSAVLTNAPYLNTLCCDHYNNSKAIKEMCFMMDPLGKCYCYQFPQ 593
DB 537 KXAGANALVAVSAVLTNAPYLNTLCCDHYNNSKAIKEMCFMMDPLGKCYCYQFPQ 596
QY 594 RPDGIDRHRDRYANRNVFFEDINMKGLDGIQGPITYGTCVFRQALYGVDAKPKKPSR 653
DB 597 RPDGIDRHRDRYANRNVFFEDINMKGLDGIQGPITYGTCVFRQALYGVDAKPKKPSR 656
QY 654 TCNCRKRCVCCCCPGRKCKTKTKTEKKLLFKKEENQSPAYALGEIDE--AARQA 711
DB 657 TCNCRKRCVCCCCPGRKCKTKTKTEKKLLFKKEENQSPAYALGEIDE--AARQA 710
QY 712 ENKAGIVNQOKLEKFKGSSVFTSTLNGSTLKSASPLSKAIHIVISGVEDKTD 771
DB 711 NVQSTBANOMKLEKFKGSSVFTSTLNGSTLKSASPLSKAIHIVISGVEDKTD 770
QY 772 WGEIGWYGSVTEEDILTGFKMCHGWSIYCIKRAVAFKSAPLMLSDRLHGVLMALG 831
DB 771 WGEIGWYGSVTEEDILTGFKMCHGWSIYCIKRAVAFKSAPLMLSDRLHGVLMALG 830
QY 832 SIIFSNHGLMVGCGGKLFERESYNSIYPMWISPLAVCTLPALCLTGKFTIP 891
DB 831 SIIFSNHGLMVGCGGKLFERESYNSIYPMWISPLAVCTLPALCLTGKFTIP 890
QY 892 ELNNVASLWMSLFCIFATSIEMRWSGVGIDMMRNEQFNVIGVSGSHLFAVFOQLL 951
DB 891 ELNNVASLWMSLFCIFATSIEMRWSGVGIDMMRNEQFNVIGVSGSHLFAVFOQLL 950
QY 952 VIAGVDTSTVTSKGGDEEFSLEYTFKWTLLIPTTLLLLNFIGVAGVSNAINNGYE 1011
DB 951 VIAGVDTSTVTSKGGDEEFSLEYTFKWTLLIPTTLLLLNFIGVAGVSNAINNGYE 1010
QY 1012 SNGPLFGKLFPAFVAVYVHLVPLKGLVGRONRTPTIYVWSIILASIFSLMWRIDPFLA 1071
DB 1011 SNGPLFGKLFPAFVAVYVHLVPLKGLVGRONRTPTIYVWSIILASIFSLMWRIDPFLA 1070
QY 1072 KDDGPLEEGGLDC 1085
DB 1071 K-GPILFICGLDC 1083

```

```

RESULT 7
Q9SJ22 PRELIMINARY; PRT: 1088 AA.
AC Q9SJ22;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Putative cellulose synthase catalytic subunit.
GN ATG21770.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE:20083487, PubMed:10617197;
RA Lin X., Kaul S., Rounsley S.D., Snea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007019; AAD20396.1; -.
DR InterPro: IPR005150; Cellulose_synth.
DR InterPro: IPR001811; Znf_ring.
DR Pfam: PF03552; Cellulose_synth. 1.
DR SMART: SM00184; RING. 1.
SQ SEQUENCE 1088 AA; 123446 MW; 060571118600DC9F CRC64;

```

```

Query Match 71.1%; Score 4209.5; DB 10; Length 1088;
Best Local Similarity 70.2%; Pred. No. 0;
Matches 769; Conservative 127; Mismatches 179; Indels 21; Gaps 10;

```

```

QY 1 MEASAGLVAGSHNRRELVTIRRDGPPKPPPEONGQVQICGDDVGLAPGDPVACNE 60
DB 1 MNTGRLVAGSHNRRELVTIRRDGPPKPPPEONGQVQICGDDVGLAPGDPVACNE 60
QY 61 CAPPCRCDCYEYERBGTGONCPOCKTRYRLKGCORVTGDEEEDGVDIDNEFNMGHDS 120
DB 61 CAPPCRCDCYEYERBGTGONCPOCKTRYRLKGCORVTGDEEEDGVDIDNEFNMGHDS 118
QY 121 GVAASML-YGMAVSGRGDPPGAPQAPQALNPNVLTGQVAVDDIPEQHALVSPFGG 179
DB 119 EVATEALVYMLNLTGRGIDVSHLYSASPGSEVPLTYCDESDSYSDRHALLIVPSTG 178
QY 180 GSKRIHPLPYADPSLPVOPRSMDSKDLAAYGSGVAMKERMENKORQ-ERM---HOT 234
DB 179 LGNRVHNVFTDSFASIHRRPVPQKDLVYGVGVAMKDRHWKKQOIELQVYKNER 238
QY 235 GNDGCG----DGDADILPLMDEARQOLSRKIPLPSSQINDPYMIIIRLVLFEPHY 289
DB 239 VNDGDGDFIVDELDPGLPMMDDEGRQPLSRKLPTRSRINPYRMILFCRLAILGLFHY 298
QY 290 RVNHPRNDAFALMISVIGCEWFMAMWILDOFPKMPRIERTYLDLSLFDKSGQPSQAL 349
DB 299 RILHPVNDAFGLMISVIGCEWFMAMWILDOFPKMPRIERTYLDLSLFDKSGQPSQAL 358
QY 350 APIDPEVSTVDPLEPPLVNTNTVLSLSVDYPRVDKSCVSDGAMLTFFALSTSESP 409
DB 359 APIDPEVSTVDPLEPPLVNTNTVLSLSVDYPRVDKSCVSDGAMLTFFALSTSESP 418
QY 410 AKKWPFCRRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVA 469
DB 419 AKKWPFCRRYNIERPAPEWYFOOKIDYLDKQVAAANFVRERRAMKREYEEFKVYNALVA 478
QY 470 KXQVPEEGMTODGTPWPGNNVVRDHPGMIOVFLGSGGLDCEGNEPLPLVYVSRKRRG 529
DB 479 VQKQVPEEGMTODGTPWPGNNVVRDHPGMIOVFLGSGGLDCEGNEPLPLVYVSRKRRG 538
QY 530 YNHKKAGAMNALVAVSAVLTNAPYLNTLCCDHYNNSKAIKEMCFMMDPLGKCYCYQFP 589
DB 539 FDKHKKAGAMNALVAVSAVLTNAPYLNTLCCDHYNNSKAIKEMCFMMDPLGKCYCYQFP 598
QY 590 QFPQRFDGIDRHRDRYANRNVFFEDINMKGLDGIQGPITYGTCVFRQALYGVDAKPKK 649
DB 599 QFPQRFDGIDRHRDRYANRNVFFEDINMKGLDGIQGPITYGTCVFRQALYGVDAKPKK 658

```

Qy	650	PSPRCNQMPKMCFFCCCGNRRKOKTKTKPTKEKKLLFFKKEENQSPAYALGEIDEAP	709
Db	659	PPKRTCNQMPKMC - CLCC - GMRK - KKTGVXDNRK --- - KPETSQKIHAEHIEBGLQ	711
Qy	710	GAENEKAGIVNQOKLEKKFFGQSSVFTSTLLENGTLLKSASBASILKKAHIVISCCGYBDK	769
Db	712	VTNANNSETAQLKLEKKFFGQSPVLVASTILLNGVSPNNVPASILLRISIQVISCYBEK	771
Qy	770	TDGKEIEMWYIGSVTEDDLITGFQKCHGMRISYICPKXVAFKGSAPLWISDRLOVLEMA	829
Db	772	TEWGEBIMWYIGSVTEDDLITGFQKCHGMRISYICMPKKAAPGASAPIMLSRLOVLEMA	831
Qy	830	IGSIEIPFSNHCPLWYGYGGGLKPLERFSYINSIYPMWTSIPILAYCTLPALCULTGKFI	889
Db	832	IGSVETIPFSRHCPIWYGYGGGLKMLERFSYINSIVPMWTSIPILAYCCLPALCULTGKFI	891
Qy	890	TPELNVASLWFMFSIFICIFATSILEMWVGVIDDMWRNEQFWYIGGVSSHLEPAVFOGL	949
Db	892	VPEISNVAGIIFLLFMFSIAVATGIIEMQMGKIGIDMWRNEQFWYIGGVSSHLEPAVFOGL	951
Qy	950	LKVIAGVQTSPLVYNSKGGDDEFSFLYFPKMTLLIIPPTLLILNFIGVAVGSNAING	1009
Db	952	LKVILAGVSTNFIYVSKAADDEFSFLYFPKMTSLIIPPTLLIINIYVAVIGVGSNAING	1011
Qy	1010	YESWGPLLEGLJFAFVWIVLHLYPLKGLVGRNRPITIVYVNSILIASIFSLIWLVRIDPF	1063
Db	1012	YDSWGPLFGRLFFFAWIVLHLYPLKGLIGKQDRVPTIILWVSILIASIITLLWVRVNP	1071
Qy	1070	LAKDGPPLLEBGLDC 1085	
Db	1072	VSK - DGPVLEICGLDC 1086	

## RESULT 8

AD	Q9F1B9	PRELIMINARY;	PRT; 1069 AA.
IC	Q9F1B9		
DT	01-MAR-2001 (TREMBLrel. 15, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE	Cellulose synthase catalytic subunit.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eustroids II; Brassicales; Brassicaceae; Arabidopsis.		
KN	NCBI_Taxid=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Columbia;		
RX	MEDLINE=99156233; PubMed=1004848;		
RA	Asamizu E., Sato S., Kaneo T., Nakamura Y., Kotani H., Miyajima N.,		
RA	Tabata S.)		
RT	Structural analysis of Arabidopsis thaliana chromosome 5. VIII.		
RT	Sequence features of the regions of 1,081,958 bp covered by seventeen		
RT	physically assigned P1 and TAC clones."		
RL	DNA Rep. 5:379-391(1998).		
DR	EMBL; AB016893; BAB09408.1; -		
DR	InterPro; IPR005150; Cellulose_synt.		
DR	InterPro; IPR001841; Znf_ring.		
DR	Pfam; PF03552; Cellulose_synt. 1.		
DR	SMART; SMO0184; RING. 1.		
DR	PROSITE; PS50089; ZF_RING_2; 1.		
SO	SEQUENCE 1069 AA; 120861 MW; D03ED5C578DB3E7C CRC64;		

## Query Match

Query Match	71.0%;	Score 4204.5;	DB 10;	Length 1069;
Best Local Similarity	70.6%;	Pred. No. 0;		
Matches 770;	Conservative 166;	Mismatches 166;	Indels 29;	Gaps 10;

Qy 1 MEASAGLVAGSHNRRELVTIRRDGDDPPRPPEQNVQCI GSDVGLAPGGDPPVACNE 60

Db 1 MNTGGRLIAGSHNRPEFVLINADESARINSVEHLSGQTQCI GSDDELVSVDGESFVACNE 60

Qy 61 CAFPLVCRCDCYEYERREGTONCPQCKTRRYRKAKCCQAVTGDDEEDGVDDLDEFNMWGHDS 120

Db	61	CAFYVCPACIYERREBNQSCPOCKITRYKIKGSPRVEGDEEDGIDDLDEFFY----	S	116
Qy	121	QSVASMLYG--HNSYGRGDPNGAPAPAFQJLNPVPLTJNQWVDIIPRGOALVPSEMG	178	
Db	117	RSGESESTFHSRNESEFLDASAPPS-----QIPLLTYEEDVEIISSDHALIVSSP	168	
Qy	179	GCGKRHPPLPADSLPVQPRSMDSKDLAAYGYSVAWKEMWKKORQRMHOT--G	235	
Db	169	GHIRVHPHPDPA--AHPRPWPQKDLAYAGYSVAMKORMEEMKKQKNEKYQVVGH	226	
Qy	236	NDGGGDDDDADLPLMDEARQQLSRKPLPESQJNPBMIIIFLVLVGFPHYRMHPV	295	
Db	227	GDSSLGDGDDIDIPMBEGROPRLSRKVPYSSKJINPRLMLVLVLVLGLFPHYRILHPV	286	
Qy	296	NDAPALMLISYICGIWPMWSWILDQPPKWPPIERETYLDRLSLFPDXGQPSOLAPIDF	355	
Db	287	NDAYALMLISYICGIWPMWSWILDQPPKWPPIERETYLDRLSLFPDXGQPSOLAPIDF	346	
Qy	356	VSTYDPLKEPPLVNTNTVLSLSDYVYDVKSCVYSDGGAAMLFELSESTSEPAKKVP	415	
Db	347	VSTYDPMKEPPLIRANTVLSLSDYVYDVKSCVYSDGGAAMLFELSESTSEPAKKVP	406	
Qy	416	FCRKATNIEPRAPENYFOOKIDYLDKVAANVREERRAMKREYEFKVRINLVAKQKVP	475	
Db	407	FCRKATNIEPRAPENYFCHKNDYLNKQKHAPAVRRRANKRDYEEFKKINLVATQKVP	466	
Qy	476	EEGNTMODGTWPGNNVADHPGMIQVFLQSGGLDCBGNBLPRLVYVRSREKRPQYNHKK	535	
Db	467	EEGNTMODGTWPGNNVADHPGMIQVFLGNNGVADVENNELPRLVYVRSREKRPQYDHHK	526	
Qy	536	AGANNALVRSAYVLTNAPYLLNDCDHYINNSKAIKEAMCMMPDLGKKVCYQPORF	595	
Db	527	AGANNALVRSAYVLTNAPYLLNDCDHYINNSKALREAMCMMPDQSGKKICYQOPORF	586	
Qy	596	DGIDRHRVYARNVVFEDINMKGDGIGQPIYVGTGCVFRQALYGYDAPKTKPSPSTC	655	
Db	587	DGIDKSDRYSARNVVFEDINMKGDGIGQPIYVGTGCVFRQALYGYDAPKTKTKRATC	646	
Qy	656	NCPMPCWCCCCFGRNRKQKTKTKPKTEKKLLFKKEBENOSPAYALGEIDBAAPAE--E	714	
Db	647	NCPMPCWCCCCFGRNRKQKTKTKPKTEKKLLFKKEBENOSPAYALGEIDBAAPAE--E	698	
Qy	715	KAGIVNOQKLEKKRGQSSVPTSTLNLNGCLTKSASPASLSLKEALINYSCEYEDKTMGK	774	
Db	699	KSPBAQDQKLEKKRGQSSVPTSTLNLNGCLTKSASPASLSLKEALINYSCEYEDKTMGK	758	
Qy	775	EIGMIGYSVTEIDITGFMHCHGRSIIYCI PKRVAFGSAPLNSDRHLQVLRNALGIE	834	
Db	759	EIGMIGYSVTEIDITGFMHCHGRSIIYCI PKRVAFGSAPLNSDRHLQVLRNALGIVE	818	
Qy	835	IFPSNHCPLWGYGGGLKFLERFSYINSIYVPMSTIPPLAYCTLPAICLLTGKPTTBELN	894	
Db	819	IFLRRHCGIWWGYGGGLKFLERFSYINSIYVPMSTIPPLAYCTLPAICLLTGKPTTBELN	878	
Qy	895	NVASLMPMSLPICFATSIITEMRBSGCIIDMMRNEOPWVIGYSSHLPAFPOGLAKYIA	954	
Db	879	NVASLMPMSLPICFATSIITEMRBSGCIIDMMRNEOPWVIGYSSHLPAFPOGLAKYIA	938	
Qy	955	GVDTSFTVTSKGGDESEBSLYTKFKWTTLLIPTTLILLNFIVGAVGSAINNGYESMG	1014	
Db	939	GVEINFTVTSKAAADGESEBSLYTKFKWTTLLIPTTLILLNFIVGAVGSAINNGYESMG	998	
Qy	1015	PLPFGKLPFAFVYVYHLIYPLKGLVGRONRTPTIVINSILIASIFSLMLWIDBPLAKD	1074	
Db	999	PLPFRLEPFAFVYVYHLIYPLKGLVGRONRTPTIVINSILIASITLMLWVYNPFVAK--G	1057	
Qy	1075	GPLLEBGLDC	1065	
Db	1058	GPLLEBGLDC	1068	

## RESULT 9

```

065338 ID 065338 PRELIMINARY; PRT; 1081 AA.
AC 065338;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Cellulose synthase (Fragment).
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Columbia;
RA Wu L., Josh C.P., Chiang V.L.;
RT "Arabidopsis, a new member of the cellulose synthase gene family from
RT Arabidopsis (Accession No. AF062485) (PGR98-114).";
RL Plant Physiol. 117:1125-1125(1998).
DR EMBL; AF062485; AAC29067.1; -.
DR InterPro; IPR005150; Cellulose_synth.
DR Pfam; PF03552; Cellulose_synth. 1.
DR SMART; SM00184; RING_1.
DR PROSITE; PSS0089; ZF_RING_2; 1.
FT NON TER 1
SQ SEQUENCE 1081 AA; 122446 MW; DC59A35A1713FD9F CRC64;

Query Match 71.0%; Score 4202; DB 10; Length 1081;
Best Local Similarity 70.8%; Pred. No. 0;
Matches 770; Conservative 126; Mismatches 172; Indels 20; Gaps 10;

QY 7 LVAGSHNRNELVIRRDGDPGKPRREGQVCOICGDVGLAFAGDPFVACNECAPVC 66
DB 4 LIAGSHNRNEVLINADENARIRSVGLSGTQICRDEIELTVDSBPVACNECAPVC 63

QY 67 RDCYEYERRBGTQNCPOCKTRRYKLGCGRTGEEEDGVDLDERNFM--DGHDSQVA 124
DB 64 RPYCYEYERRBGTQNCPOCKTRRYKLGCGRTGEEEDGVDLDERNFM--DGHDSQVA 123

QY 125 ESMLYGHMSYG--RGDNGAQAQAFQALNPVPLTNGQVNDIPPEQHAL--VPSFMGGGK 182
DB 124 ESMYISRRNSGFFPQSDSDASAPGSG---IFLLTYGDBDVEISDRHALVPSFGHGN 179

QY 183 RIHPLPYADPSLPVQPSKMPKSKDLAAVYGVSVAKRMENMKORO--ERMHQTGNDGCD 241
DB 180 RVHGVLSJSDPTVAHRRILMPQKDLAYGVGVAMKORMBEMKMKQKHEKQVNVHEDPD 239

QY 242 --DGGDADPLMDEARQOLSKRILPSSQINPYMIIIRLVLGFFPHTRVMAFVNDAF 299
DB 240 FEDGDDADPFMMDEGRQPLSKWIKIPKSKINPYMILVLRVLIGLFFPHYRILHVPDXY 299

QY 300 ALMLISYICETWFMMSWITLOPFMPFPIERTETVLDLSTIRDKGCSQALPIDFVSTV 359
DB 300 ALMLISYICETWFMMSWITLOPFMPFPIERTETVLDLSTIRDKGCSQALPIDFVSTV 359

QY 360 DPLKEPPLVNTNTLSISLVDPYDKVSCVYSDGAMLTLEALSEFSEFAKKNVPCCKR 419
DB 360 DPLKEPPLVNTNTLSISLVDPYDKVSCVYSDGAMLTLEALSEFSEFAKKNVPCCKR 419

QY 420 YNIEPRAPWYFOQKIDYLDKQVAVNFRERRANKREYEEFKVIRINLVAKQVPEEGW 479
DB 420 YNIEPRAPWYFOQKIDYLDKQVAVNFRERRANKREYEEFKVIRINLVAKQVPEEGW 479

QY 480 TMOGCTWPGNSVDHGMIOVPLFGSGGLDCEGNELPRLVYVREKRPQGNHKKKXGAM 539
DB 480 TMOGCTWPGNSVDHGMIOVPLFGSGGLDCEGNELPRLVYVREKRPQGNHKKKXGAM 539

QY 540 NALRVASVLTNAEYLLNTLDCDHYINNKAKEAMCFMMDPLKSKVYVQFPQRFQID 599
DB 540 NSLIRVSGVLSNAPYLLNVDCDHYINNKAKEAMCFMMDPQSGKICYVQFPQRFQID 599

QY 600 RHDYANRNVVFPFIINKGDLGIQGPITYGTGCVFRQALYGYDAPKTKKPSRTCNQWP 659

```

```

DB 600 RHDYANRNVVFPFIINKGDLGIQGPITYGTGCVFRQALYGYDAPKTKKPSRTCNQWP 659
QY 660 KMFCCCCCGNRRKOKKTTKPKTEKKKLLFFKKEENOSPAYALGEIDEAA--DGAENKAG 717
DB 660 KMWL--LCGSRNRKRAKTVVADKK---KREASQKHALENIEGSHVLANVEQST 713
QY 718 IYNQQLKLEKFGGSSVFTSTLLENGCTLKSAPSLKEAIVHISGVEDTDMGKEIG 777
DB 714 EAMQWLQKRYGSPFVVASRLENGAMRNASPACLLEAIOVSRGYEDKTEWKEIG 773
QY 778 WIVGSYTEDILTGFKNGHGWRSIYCIKRVVAFKGSAPLNSDRLHQVRNALGSEIEFP 837
DB 774 WIVGSYTEDILTGFKNGHGWRSIYCIKRVVAFKGSAPLNSDRLHQVRNALGSEIEFP 833
QY 838 SNHCPLMYGGGLKELERFSYINSIVYPMTSIPLLAYCTLPALCLTGKFTPBELANYA 897
DB 834 SRRCPLMYGGGLKELERFSYINSIVYPMTSIPLLAYCTLPALCLTGKFTPBELANYA 893
QY 898 SLTFMSLFICIFATSLLENRWSGVGIDMWRNEQFVIGVSSHLPAVFOGLIKVTAGVD 957
DB 894 SLTFMSLFICIFATSLLENRWSGVGIDMWRNEQFVIGVSSHLPAVFOGLIKVTAGVD 953
QY 958 TSFTVTSKGDDEESELTFKMTTLLIPTLLILNPIGVAAGVSNATNNGVESGRLP 1017
DB 954 TSFTVTSKADDEEFDLTFKMTSLIPMTLLINPIGVAAGVSNATNNGVESGRLP 1013
QY 1018 GKLFPAFWYVHLHYPLKGLVGRONTPTIVVMSILASISLAVRIDPFLAKDQDGL 1077
DB 1014 GRFLPFWYVHLHYPLKGLVGRONTPTIVVMSILASISLAVRIDPFLAKDQDGL 1072
QY 1078 LEECGLDG 1085
DB 1073 LEICGLDG 1080

RESULT 10
Q9L19 ID Q9L19 PRELIMINARY; PRT; 1075 AA.
AC Q9L19;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cellulose synthase-1.
GN CESA-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PNCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xocoostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
RT family."
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200525; AAF89861.1; -.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF03552; Cellulose_synth. 1.
DR SMART; SM00184; RING_1.
SQ SEQUENCE 1075 AA; 121181 MW; 67A4FBD97A811F33 CRC64;

Query Match 67.3%; Score 3984; DB 10; Length 1075;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 738; Conservative 143; Mismatches 167; Indels 64; Gaps 15;

QY 1 MEASAGLVAGSHNRNELVIRRDG--PG--PKPREQNGQVCOICGDVGLAFAGDPFVAC 58
DB 1 MAANKGMVAGSHNRNEFVMI RHDGDPGSAKPKTSANGVCOICGDSVGVATGDFVAC 60

```



QY 577 NMDDPLGKKVYVQPPQFQFDGIDRDRYANRVNVEFDINMKGLDGIQGIYVGTGCVERR 636  
 DB 574 NMDDPALGRKTCYVQFPQFQFDGIDLDRYANRVNVEFDINMKGLDGIQGVVGTGCVERR 633  
 QY 633 QALYGDARKTKKPPBRTCNCPKPKFCFCCCGNKKQKTKTPKTEKKLLFPKKEENS 696  
 DB 634 QALYGD-----PVLTEADLEPNIVVKS--GRRRRKKKSYVDSGR---IMKRTSSA 683  
 QY 697 PAVALGEIDEAAPGAENEKAGIVNOOKLEKKFGQSSVFVSTSTLLENGTGLKASPAASLTK 756  
 DB 684 PFNNMEDIEEGIEGYDERSVLMQKLEKRGQSIPTASTFMQGGIPSTNPASLTK 743  
 QY 757 EAHIVISCGYEDKTDWKEIGWYGSVTEBIIITGFQMGHGRSIYCIKRYAFKGSAPL 816  
 DB 744 EAHIVISCGYEDKTDWKEIGWYGSVTEBIIITGFQMGHGRQSIYCMPPRCFKGSADI 803  
 QY 817 NISDRLOVLRWALGSIETIFPSNHCPLMTGYGGGLKFLERFSYINSIYVPMISIPLANC 876  
 DB 804 NISDRLOVLRWALGSIETILSRHCPIMWGYNGRLRLRLAYINTIVPTISVPLIAYC 863  
 QY 877 TLPAICLLTGKFTPELNNVASLPMFSLFICIPATSIILSMRSGVIGIDDMRNROFPWIG 936  
 DB 864 VLPALICLLTKKFTPELNNVASLPMFSLFICIPATSIILSMRSGVIGIDDMRNROFPWIG 923  
 QY 937 GVSSHLFAVFOGILLKVIAGVDTSPVTSKGD-DEEFSSELYTFKWTLLIPPTLLLNLF 995  
 DB 924 GTSANLFAVFOGILLKVIAGVDTSPVTSKGD-DEEFSSELYTFKWTLLIPPTLLLNLF 983  
 QY 996 IGVVACVSNAINNGVSKGPFGLKFAFMVIVLHLYPLKGLVGRONRPTIYVWSILL 1055  
 DB 984 VGVVACVSNAINNGVSKGPFGLKFAFMVIVLHLYPLKGLVGRONRPTIYVWSILL 1043  
 QY 1056 ASIFSLWVRIDPEFLA-KDDGPILEECGJDC 1085  
 DB 1044 ASIFSLWVRIDPEFLA-KDDGPILEECGJDC 1074  
 RESULT 12  
 Q9SMW6 PRELIMINARY; PRT; 1026 AA.  
 ID Q9SMW6  
 AC 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DE 01-MAY-2003 (TREMblrel. 23, Last annotation update)  
 GN IRX3 OR T10B6.80  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eucoside II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_Taxid=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANDSBERG ERECTA;  
 RX MEDLIN=99264300; PubMed=10330464;  
 RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;  
 RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose  
 RL synthase required for secondary cell wall synthesis.";  
 RL Plant Cell 11:769-780(1999).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,  
 RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Meyer K.F.X.;  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RA Shum P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,  
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,  
 RA Heuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,  
 RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
 RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis cDNA clones";  
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF091713; AAD40885.1; -  
 DR EMBL; AL391142; CAC01737.1; -  
 DR EMBL; AY139754; AAM98075.1; -  
 DR InterPro; IPR005150; Cellulose\_sync.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF03552; Cellulose\_sync; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 SQ SEQUENCE 1026 AA; 115797 MW; 503BFC78B6E511 CRC64;  
 Query Match 67.0%; Score 3967.5; DB 10; Length 1026;  
 Beet Local Similarity 66.8%; Pred. No. 0;  
 Matches 736; Conservative 129; Mismatches 145; Indels 91; Gaps 16;  
 QY 1 MEASAGLVAGSHNRRELTVIRRDGPPGPPREQNGVQICDGVGLAPGDPFVACNE 60  
 DB 1 MEASAGLVAGSHNRRELTVIRNHEE--PKPLKNDGQFCETICGDIIGLTVBGLFVACNE 58  
 QY 61 CAPVYCRDCYERREBGTQNCPOCKTRYKRLKGCQRTVDEEDGVDDLNEFMWGHDS 120  
 DB 59 CGFPACRDCYERREBGTQNCPOCKTRYKRLKGRSPPRVGDEDEIDIDIEYFNIE--HEQ 117  
 QY 121 ---GSVAESMLYGHMSYRG--GDPNGA--PQAFQANPVPLTLNGQWVDOLPPEQHALLP 174  
 DB 118 DKHKSAAEMLYGKMSYRGEDENGREF-----PVIAGHSGE----- 157  
 QY 175 SFMGSG-----KRILPLYPADPLPVPQPSMDPSKDLAAYGYSVAMKERENMK 225  
 DB 158 -PVPVGGYNGENHGLKRVHP-----PSSASGSG--GKREMDMK 197  
 QY 226 QROERHQTGNDGGDDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVIGF 285  
 DB 198 L-----OHGNIQPEPD--DDEPMGLIDEARQPLSRKVPVIASSKIPYRMVIVARLVILAV 250  
 QY 286 PFHYRVHNPVNDAPFALMLISVCEIWFAMSWITLQFPKPFPERETVLDRLSRPDKSG 345  
 DB 251 FRYRYLNPVNDALMLTSLVCEIWFAMSWITLQFPKPFPERETVLDRLSRPDKSG 310  
 QY 346 PSQALPDIFFVSTVDPPLKEPLVTNTVLSILSVDPYDKVSCYVSDGAAMLTREALSE 405  
 DB 311 PMLMLPVDVSTVDPPLKEPLVTNTVLSILSVDPYDKVSCYVSDGAAMLTREALSE 370  
 QY 406 TSEFAKKVPPCKRYNIEPRAPEMYFOQKIDYLDKQVIANPVRERRAMKREVEEFKVIN 465  
 DB 371 TAEFARKVPPCKKFSIEPRAPEMYFTLVLDYLDKQVIANPVRERRAMKREVEEFKVIN 430  
 QY 466 ALVAKAQVPEBGMQOGTIPRGNNVADHGMIOVFQSGGGLDCBENELPRLVYVRE 525  
 DB 431 AQVAKASVPLEGWMQOGTIPRGNNVADHGMIOVFQSGGGLDCBENELPRLVYVRE 490  
 QY 526 KAPGNNHKKAGAMALRVSAVLTNAPVLTNDCDHYINNSKATKAMCFMWDPLGLK 585  
 DB 491 KAPGNNHKKAGAMALRVSAVLTNAPVLTNDCDHYINNSKATKAMCFMWDPLGLK 550  
 QY 586 VCYVQFPQFQFDGIDRDRYANRVNVEFDINMKGLDGIQGIYVGTGCVRRQALYGDAP 645  
 DB 551 VCYVQFPQFQFDGIDRDRYANRVNVEFDINMKGLDGIQGPVYVGTGCVRRQALYGERP 610  
 QY 646 KTKKPPSRTCNCWPMKPCFCCCFGRKQKTKTKPKTEKKLLFPKKEENQSAVALGEID 705  
 DB 611 KTKKPPSRTCNCWPMKPCFCCCFGRKQKTKTKPKTEKKLLFPKKEENQSAVALGEID 649  
 QY 706 EAPGAENEKAGIVNOOKLEKKFGQSSVFVSTSTLLENGTGLKASPAASLTKAHHVISC 765  
 DB 650 ----GAEQCKHLMENMFEKTFGQSSIFVISTLMBEGVPPSSPAVLKKAHHVISC 705  
 QY 766 YEDKTDWKEIGWYGSVTEBIIITGFQMGHGRSIIYCIKKRVAKGSAPLNLSDRLOHV 825

```

Db 706 YEDKTEWGTGLGWIIGSTIEDILTFGFKHCRGMRSIGCMKPAKGSAPILSLRNLQV 765
Qy 826 LRMALGSIETFFSNHCPLMYG--GGGLKLELFSTINSIVPWSIPILAYCTLPALICL 884
Db 766 LRMALGSVEIFFSRHSPLMYGKGKLMKLERFAVANTTIYPFTSIPILAYCILPAICL 825
Qy 885 TGFPTPLNNVASLFWMSLFCIFATSTILEMNSGVGIDDMWRNEQFWVIGVSHLFA 944
Db 826 TDKFIMPLSTFASLFFLSLFWMSIIVTGLLELMSGVSTLEEMWRNEQFWVIGVSHLFA 885
Qy 945 VFQGLKLYIAGVDTSFYTSKGGDDEFSLEYTFKMTLLIPPTLLILNFIIVAGVSN 1004
Db 886 VVQGLKLYIAGVDTFMTYTSKATDDDFGELAFKMTLLIPPTLLILNFIIVAGVAGISD 945
Qy 1005 AINNGESWGLPFGKLPFAFWIVHLVYPLKGLVGRQRPTTIIVWSILLASIFSLMW 1064
Db 946 AINNGYQSWGLPFGKLPFSFWIVHLVYPLKGLMGRQRPTTIIVWSILLASIFSLMW 1005
Qy 1065 RIDPFLAKDDGPLLECGIDC 1085
Db 1006 RIDPFLVLTGKGPDTSKGICINC 1026

RESULT 13
Q9XHP6 PRELIMINARY; PRT, 1026 AA.
ID Q9XHP6;
AC 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DR 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN IRX3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_Taxid=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RA MEDLINE:99264300; PubMed:10330464;
RX Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
synthase required for secondary cell wall synthesis.";
RL Plant Cell 11:769-780 (1999).
DR EMBL: AF088917; AAD32031.1; -
DR InterPro: IPR005150; Cellulose_synth.
DR InterPro: IPR001841; znf_ring.
DR Pfam: PF03552; Cellulose_synth; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00089; ZF_RING_2; 1.
SQ SEQUENCE 1026 AA; 115858 MW; 453BFD1D283C4D70 CRC64;

Query Match 67.0%; Score 3963.5; DB 10; Length 1026;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 736; Conservative 128; Mismatches 146; Indels 91; Gaps 16;

```

```

Qy 226 QROBRMOTGNDGGDDADLPLMDBARQOLSRKIPLPSSQINPYRMIIIRLVVYGF 285
Db 198 L-----QHGILGSEPP--DDEPMGLIDEARQPLSRKPIASCKINPRMIVVALVILAV 250
Qy 286 FFHYRVHVPVNDAPALMLISYCEIWFAMSWILDQFPKMPPIREBYDLRLSLRFDKQ 345
Db 251 FLVRILNPVHADGLWLTSTYCEIWFAMSWILDQFPKMPPIREBYDLRLSLRBERGE 310
Qy 346 PSQAPIDFSTYDPLKPEPLVTTNTVSLSTVDYPRDKSCVSDGAAMLTPEALSE 405
Db 311 PNMILAPVDVFTSTYDPLKPEPLVTSNTVSLSTAMDYPERKISCVSDGAAMLTPEALSE 370
Qy 406 TSEPAKKVPCKKYNIEPRAPENYFOQKIDYLDKXAAANVRBRRAKRESEPKVRIN 465
Db 371 TABFARKVVPCKKFSLEPRAPENYFTLKVDYLDCKHAPTYKERRAKRESEPKVRIN 430
Qy 466 ALVAAQKVPPEGTMQDTPMPGNVNDHFGMIQVFLGQSGDLDEGNEPLRLVYVRSRE 525
Db 431 AQVAKASVPLEGIMQDTPMPGNNTKDHFGMIQVFLGHSQGFVDEGHELPRLYVRSRE 490
Qy 526 KRPGYNHKKAGANALVRVASAVLTNAPEYLLNDCDHYNNSKAIKEAMCEMDPLGKK 585
Db 491 KRPGFQHNKKAGANALVRVAVLTNAPEYLLNDCDHYNNSKAVREAMCEMDPLGCK 550
Qy 586 VCVYQPPQRFQIDRHRBYANRVFPDINNGKIDGIGPIYNGTGCYFRQALYGNAP 645
Db 551 VCVYQPPQRFQIDNDRYANRVFPDINNGKIDGIGPYVYVGVCFKQALYGNAP 610
Qy 646 KTKRPPSPTNCWMPKCFCCCFGFRKOKTKTKTEKKULLPFKGENQSPAYALGEID 705
Db 611 KGRPRPKMIS-----GCCPCFGRKRNKK-----FSKIDMDGDVAALG--- 649
Qy 706 BAAPAEANEKAGIVNQQLEKFKGQSSVFTSTILLENGTILKASAPSLKEALHVISCG 765
Db 650 ---GABGDKHLMFENMFEXTFGQSSIFVTSTLMEBGVPPSSPAVLKEALHVISCG 705
Qy 766 YEDKTEWGTGLGWIIGSTIEDILTFGFKHCRGMRSIGCMKPAKGSAPILSLRNLQV 765
Db 706 YEDKTEWGTGLGWIIGSTIEDILTFGFKHCRGMRSIGCMKPAKGSAPILSLRNLQV 765
Qy 826 LRMALGSIETFFSNHCPLMYG--GGGLKLELFSTINSIVPWSIPILAYCTLPALICL 884
Db 766 LRMALGSVEIFFSRHSPLMYGKGKLMKLERFAVANTTIYPFTSIPILAYCILPAICL 825
Qy 885 TGFPTPLNNVASLFWMSLFCIFATSTILEMNSGVGIDDMWRNEQFWVIGVSHLFA 944
Db 826 TDKFIMPLSTFASLFFLSLFWMSIIVTGLLELMSGVSTLEEMWRNEQFWVIGVSHLFA 885
Qy 945 VFQGLKLYIAGVDTSFYTSKGGDDEFSLEYTFKMTLLIPPTLLILNFIIVAGVSN 1004
Db 886 VVQGLKLYIAGVDTFMTYTSKATDDDFGELAFKMTLLIPPTLLILNFIIVAGVAGISD 945
Qy 1005 AINNGESWGLPFGKLPFAFWIVHLVYPLKGLVGRQRPTTIIVWSILLASIFSLMW 1064
Db 946 AINNGYQSWGLPFGKLPFSFWIVHLVYPLKGLMGRQRPTTIIVWSILLASIFSLMW 1005
Qy 1065 RIDPFLAKDDGPLLECGIDC 1085
Db 1006 RIDPFLVLTGKGPDTSKGICINC 1026

RESULT 14
Q9GSM2 PRELIMINARY; PRT, 1032 AA.
ID Q9GSM2;
AC 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DR 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Cellulose synthase.
GN CES42.
OS Populus tremuloides (Quaking aspen).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```



OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucoside 1; Malpighiales; Salicaceae; Populus.  
 OC NCBI\_TaxID=3693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Xylem;  
 RX MEDLINE=22271539; PubMed=1283501;  
 RA Samuga A.; Joshi C.P.;  
 RT "A new cellulose synthase gene (PtxCea2) from aspen xylem is  
 RT orthologous to Arabidopsis AtCea7 (Atx3) gene associated with  
 RT secondary cell wall synthesis.";  
 RL Gene 296:37-44(2002).  
 DR EMBL:AY095297; AAC26299.1; -  
 SQ SEQUENCE 1032 AA; 116958 MW; 15ABB28FDC60B96D CRC64;  
 Query Match 66.4%; Score 3955.5; DB 10; Length 1032;  
 Best Local Similarity 66.4%; Pred. No. 0;  
 Matches 729; Conservative 138; Mismatches 152; Indels 79; Gaps 15;  
 QY 1 MEASAGLVAGSHNRNELVIRRDGDPKPRPREGQVQCIGDDVGLAPGGDPFYACNE 60  
 DB 1 MEASAGLVAGSHNRNELVIRH--GHEHNRPLKMLDQVCEICGDEIGLTVDDGLFYACNE 58  
 QY 61 CAPFVRCDCYERREBGTGNCPOCKTRYRKLGCRCQVTDDEEDGVDDLDNERNDGHDS 120  
 DB 59 CGFVRCPCYERREBGTGNCPOCKTRYRKLGCRCQVTDDEEDGVDDLDNERNDGHDS 118  
 QY 121 QS--VAESMLYGHMSYRGDPPGAPQAFQALNPVPLTLTGQVAVDDIPEQHALVPSFNG 178  
 DB 119 KKHHLTEAMLHGMYTGRGHD-----EENSQFPVLTGIRSRVSGE-----FSIG 165  
 QY 179 GGG-----KRHPLPYADPSLPVQPSMDPSDGLAAYGYGVAMKERNMKQROE 229  
 DB 166 SHGEQMLSSLSHKRVAPVSESPG---SARWDEKKE-----GGMKERMDWEMK--- 210  
 QY 230 RHHQGNDDGGDGDADLPLMEARQOLSRKILPSSQINPYRMIIIRLVLTGFFHY 289  
 DB 211 ---QHGNLDPEDD--DAEAMLEDAQPLSRKVLASSKINPRMYIVARLIIILAVFLRY 266  
 QY 290 RVHVPNDAPFALMLISVCEIFWPMWMIIDQFPKFPPIRETYLDLSLRFDEGQPSOL 349  
 DB 267 RILHPRHDLGLMLTSIVECEIWMIDQFPKMLPIRETYLDLSLRFDEGQPSOL 326  
 QY 350 APIDFVSTVDPLKEPPLVTTNTVLSLVDPVVDKVCVSDGAMLTFEALSTSEF 409  
 DB 327 APVDVSTVDPLKEPPLVTTNTVLSLVDPVVDKVCVSDGAMLTFEALSTSEF 386  
 QY 410 AKKVPFCRKYNERPAPRYFOOKIDYLDKVAAPFVERBRMKREYEFKRIATAVA 469  
 DB 387 ARKVPFCRKYNERPAPRYFOOKIDYLDKVAAPFVERBRMKREYEFKRIATAVA 446  
 QY 470 KAKVPEEGWTMDGTPWPGNNVDRHGMIOVFLGSGGLDCEGNEPLRLVYVSRKRG 529  
 DB 447 KAKVPEEGWTMDGTPWPGNNVDRHGMIOVFLGSGGLDCEGNEPLRLVYVSRKRG 506  
 QY 530 YNHKKAGAMNALVRSVAVLTNAAYLLNDCHYINNSAKIEMKFMMDPLIGKVCYV 589  
 DB 507 FSHKKAGAMNALVRSVAVLTNAAYLLNDCHYINNSAKIEMKFMMDPLIGKVCYV 566  
 QY 590 QPQORPDGIDRHRYANRVNVPEDIMKGLDGIQGIYVGTGVFROALYGDARK-7K 648  
 DB 567 QPQORPDGIDRHRYANRVNVPEDIMKGLDGIQGIYVGTGVFROALYGDARK-7K 626  
 QY 649 KPSRTQNCWPKMFCFCCGNGRKKOKTKTKTEKKKLTFKKEENSPAYALGEIDEAA 708  
 DB 627 RPKMEKCD-----CCPCGRKKKNA---KT-----GAVVEGM 656  
 QY 709 PGANEKAGIVNQQLEKKFGSSVPTSTLLENGTASASAPSLKEAIIHISCGYED 768  
 DB 657 D--NNKELMASHNNEKKFGQSAIFATSTLMEEGVPPSSPAALLKEAIIHISCGYED 714  
 QY 769 KTDGKEIGWYGVSVEDILTGFMCHGGRSTYCI PKRAAFKGSAPLNSDRLHVLWM 828  
 DB 769 KTDGKEIGWYGVSVEDILTGFMCHGGRSTYCI PKRAAFKGSAPLNSDRLHVLWM 828

DB 715 KTEMGLGWIYGTSTEDILTGFMCHGGRSTYCI MPGRAPFKSAPINLSDRNLQVLRM 774  
 QY 829 ALGSIIEFSSNHCPLWGY-GGGKFLERFSYINSIYVPMWISPLAVCTLPALCLTGG 887  
 DB 775 ALGSVEIFFSGHSPWYGVKKGKMLERFAVYVNTTIYFPSTLAVVCCPAILCLLTDK 834  
 QY 888 FITPELNVASLMPFSLFCIFATSILEMRSYGICIDDMWNEQFVWYIGVSSHLPAVQ 947  
 DB 835 FIMPEISFVASLFFIALFLSIFSGLIELRMSGVSIIEEMWNEQFVWYIGVSSHLPAVQ 894  
 QY 948 GLKVIAGVDSFVTSKGDDEPSELVTFKMTLLIPPTLLNLNPIGVAGVSNAIN 1007  
 DB 895 GLKVIAGVDSFVTSKATDDDFGELIYAKMTLLIPPTLLNLNPIGVAGVSDAIN 954  
 QY 1008 NGYESMGPFLGKLPFAFVIVHLYPFLKGLVGRONRPTIYVWSIILASIFSLWVRID 1067  
 DB 955 NGYQSGWGLPFGKLPFAFVIVHLYPFLKGLVGRONRPTIYVWSIILASIFSLWVRID 1014  
 QY 1068 PFLAKDDGPLEBECGLDC 1085  
 DB 1015 PFMKTRGPDTRKQCGLNC 1032  
 RESULT 15  
 ID 048946 PRELIMINARY; PRT; 1081 AA.  
 AC 048946;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Cellulose synthase catalytic subunit.  
 GN RSW1 OR FB84.110 OR AT4G32410.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucoside 1; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=9811412; PubMed=9445479;  
 RA Arioli T., Peng L., Betzner A.S., Burn J., Witcke W., Herth W.,  
 RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,  
 RA Redmond J., Williamson R.E.;  
 RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";  
 RL Science 279:717-720(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terry N., Ardiles W., Buyschaert C., Daseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Villarejo R., Gleien J., Van Montagu M., Hohnlel J., Mewes H.W.,  
 RA Mayer K.F.X., Lemcke K., Schueller C.;  
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Terry N., Ardiles W., Buyschaert C., Daseville R., De Clerck R.,  
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villarejo R.,  
 RA Gleien J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.,  
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.  
 DB EMBL:AF027172; AAC39334.1; -  
 DB EMBL:AL034567; CA42568.1; -  
 DB EMBL:AL161581; CAB79958.1; -  
 DR InterPro: IPR005150; Cellulose\_synth.  
 DR InterPro: IPR001841; Znf ring.  
 DR Pfam: PF03552; Cellulose\_synth; 1.

DR SMART: SM00184; RING: 1.  
SQ SEQUENCE 1081 AA; 122236 MM; BDBSD9DBE334D59 CRC64;

Query Match 66.7%; Score 3945.5; DB 10; Length 1081;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;

```

QY 1 MEASAGLVAGSNRRNELVIRRDGDPGPRPRRQNGVQICGDDVGLAPGDDPFACNE 60
DB 1 MEASAGLVAGSNRRNELVIRRHSDGDTPLKNNMQICQICGDDVGLAETGDDVFAACNE 60
QY 61 CAFPVCRDCYEXERREBGTQNCPOCKTRRYRLKGCQRVTGDEEBEDVDLDFNFMNDGHS 120
DB 61 CAFPVCRDCYEXERREBGTQNCPOCKTRFRHRKSPFVGBEDBDVDLDFNFMNDGHS 116
QY 121 QSVASMLYGHNSYGRGDPNGAPQAFQALNPVPLTNGQWVD--1PPEQHALVSPF- 176
DB 117 QGANXA-----RHQHGEEFSSSSRHESQ--1PILTHGHTVSGEIRTPDQSVRTSGP 169
QY 177 MCGGGRRIHPLPYADSLVQPRSDPSTDLAAYGYGSAWKERENMKOROR--MHOT 234
DB 170 LGPSDRNAISSPPIPRQPVPRVIVDPSTDLNSYGLGNVDMKEKRVGWLKQKNNLQMT 229
QY 235 GN-----DGGGDDGDDADLPLMDEARQQLSRKILPSSQINPYRMIIIRLVYLG 285
DB 230 GRYHECKGEIEGTGNGE--ELQMDADTRLPMRSRVPLPSSRLTPYRVIIIRLILCF 287
QY 286 PFHYRVHVPNDAFALMLISVICEIPFAMSMILDOCPKMPFIERETYLRLSLRFDKEG 345
DB 288 FLOYRTHVKNAYPLMLTSVCEIFAFSWLDDQFPKMYPIKRETYLRLAIRYDRGE 347
QY 346 PSQALAFIDFVSTVDEPLKEPPLVTNTVLSTLSVDYPVDKVCYVSDGGAAMLTFEALSE 405
DB 348 PSQLVAVDVSTVDEPLKEPPLVTANTVLSLVDYPVDKVCYVSDGSAAMLTFEALSE 407
QY 406 TSEFAKKWPFCKRYNIEPRAPEMVFOQKIDYLDKQVAAVFREERAMREYEERFVRIN 465
DB 408 TAEFAKKWPFCKRYNIEPRAPEMVFOQKIDYLDKQIOPSFYERAMREYEERFVRIN 467
QY 466 ALVAKAKVPEEGMTQDGTMPMGNNVRDHPGMIOVFLGOSGLDCGNEPLRLVYSRE 525
DB 468 ALVAKAKVPEEGMTQDGTMPMGNNVRDHPGMIOVFLGOSGLDCGNEPLRLVYSRE 527
QY 526 KRPYNNHKKAGAMNALVVSAYLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLLGKX 585
DB 528 KRPYNNHKKAGAMNALVVSAYLTNAYLTNDCDHYINNSKAIKEAMCFMMDPLLGKX 587
QY 586 VCIVQFPQRFDDGIDRRDRYANRVVFFDINMGKLDGICQPIYVGTGCVRRQALYGYDAP 645
DB 588 CCYVQFPQRFDDGIDRRDRYANRVVFFDINMGKLDGICQPIYVGTGCVRRQALYGYDAP 647
QY 646 KTKKKPSRTQNCMPKCFCCCGGNRKOKTKTKTEKKKLFFKKEENQSPAYALGEID 705
DB 648 LTEE-----DLEPNIIVKSCGSRKKGSSKXNTEKRG--NRSDSNAPLFNMDTD 699
QY 706 EAPGAENKAGIVNOQKLEKFGQSSVEVTSTLLENGGLKASAPASLKEAIVHISCG 765
DB 700 EAPGAENKAGIVNOQKLEKFGQSSVEVTSTLLENGGLKASAPASLKEAIVHISCG 759
QY 766 YEDKTDWKEIGIYGSVTEBILTFGRMKCHGRSTYICIPKRYAFKGSAPLNSDRLOV 825
DB 760 YEDKTDWKEIGIYGSVTEBILTFGRMKCHGRSTYICIPKRYAFKGSAPLNSDRLOV 819
QY 826 LEMALGSIIFPSNHCPLWYGGGLKFLERFINSIYPMWTSIPLAYCTLPAICLT 885
DB 820 LEMALGSIIFPSNHCPLWYGGGLKFLERFINSIYPMWTSIPLAYCTLPAICLT 879
QY 886 GKRTPELNVAASLWMSLFCIFATSIEMRWSGVGIDWENEOFWYIGVSSHFAV 945
DB 880 DRITPEISNVASIWTLFLFISIAVTGIELRWSGVSIDWENEOFWYIGVSSHFAV 939
QY 946 FQGLLVIAGVDTSTFTVTSKGD--DEEFSELYFKWTLILPTTLLLNPIGVAVGVS 1004
DB 940 FQGLLVIAGVDTSTFTVTSKATDEBDGPAELIYFKWTLILPTTLLLNPIGVAVGS 999

```

```

DB 940 FQGLLVIAGVDTSTFTVTSKATDEBDGPAELIYFKWTLILPTTLLLNPIGVAVGS 999
QY 1005 AINNGYSWGLFGLFPFAFVYVHLYPFLKGLVGRQNRTPPTIYVMSILLASIFSLMV 1064
DB 1000 AVNSGYSWGLFGLFPFAFVYVHLYPFLKGLVGRQNRTPPTIYVMSILLASIFSLMV 1059
QY 1065 RIDPFL 1070
DB 1060 RINPFV 1065

```

Search completed: December 15, 2003, 13:04:19  
Job time : 51 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:56:44 ; Search time 46 Seconds

(without alignments)  
3747.327 Million cell updates/sec

Title: US-09-720-383C-10

5918

Perfect score: 1 MEASAGLVAGSHNNELVVI.....DPLAKDDPILRECGLCDN 1086

Scoring table: BLOSUM62

GAPOP 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5918	100.0	1086	21	AA1984111
2	5918	100.0	1086	21	AA1984116
3	5918	100.0	1086	21	AA198835
4	5570.5	94.1	1148	21	AA198832
5	5402.5	91.3	1058	21	AA1984107
6	5328	88.3	1094	21	AA1984113
7	5328	88.3	1094	21	AA1984118
8	5328	88.3	1165	21	AA198834
9	4278.5	72.3	1084	23	AA198304

10	4262	72.0	1084	23	AA1983949	Herbicideally activ
11	4250.5	71.8	1084	19	AA1983818	Herbicideally activ
12	4209.5	71.1	1088	23	AA1981831	Herbicideally activ
13	4204.5	71.0	1069	23	AA1983411	Herbicideally activ
14	4185.5	70.7	1080	22	AA19837895	Herbicideally activ
15	3986	67.4	1039	21	AA198837	Soybean cellulose
16	3984	67.3	1075	21	AA1984109	Amino acid sequenc
17	3984	67.3	1075	21	AA1984115	Amino acid sequenc
18	3984	67.3	1075	21	AA1984120	Amino acid sequenc
19	3977.5	67.2	1074	21	AA1984112	Amino acid sequenc
20	3977.5	67.2	1074	21	AA1984117	Amino acid sequenc
21	3977.5	67.2	1074	21	AA1984112	Amino acid sequenc
22	3967.5	67.0	1026	23	AA1983522	Herbicideally activ
23	3945.5	66.7	1081	19	AA1983817	Herbicideally activ
24	3945.5	66.7	1081	23	AA1983199	Herbicideally activ
25	3935.5	66.5	1081	19	AA1983820	Herbicideally activ
26	3911	66.1	1076	21	AA1984121	DNA encoding a mai
27	3903.5	66.0	1039	19	AA1983309	Cellulose synthase
28	3899.5	65.9	1077	21	AA1984108	Amino acid sequenc
29	3899.5	65.9	1077	21	AA1984114	Amino acid sequenc
30	3898.5	65.9	1065	21	AA198733	Herbicideally activ
31	3898.5	65.9	1065	23	AA1983356	Herbicideally activ
32	3887.5	65.7	1054	21	AA198734	Herbicideally activ
33	3880.5	65.6	1079	23	AA1980496	Corn cellulose syn
34	3876.5	65.5	1065	19	AA1983819	Herbicideally activ
35	3804.5	64.3	1043	23	AA1983701	Herbicideally activ
36	3782	63.9	1065	23	AA1981880	Amino acid sequenc
37	3780.5	63.9	1043	21	AA1984119	Herbicideally activ
38	3603	60.9	959	21	AA198735	Soybean cellulose
39	3521	59.5	974	19	AA1983308	Cellulose synthase
40	3476.5	58.7	974	19	AA1983308	Cellulose synthase
41	3469.5	58.6	974	19	AA1983308	Cellulose synthase
42	3356	56.7	955	23	AA1981638	Eucalyptus grandis
43	3329.5	56.3	958	23	AA1982981	Herbicideally activ
44	3318	56.1	821	23	AA1980495	Corn cellulose syn
45	3138.5	53.0	946	22	AA1987893	Populus tremuloides

#### ALIGNMENTS

RESULT 1	AA1984111
ID	AA1984111 standard; Protein; 1086 AA.
XX	AA1984111;
AC	
XX	
DT	03-JUL-2000 (first entry)
XX	
DB	Amino acid sequence of a maize cellulose synthase.
XX	
KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW	transgenic plant; plant breeding marker.
XX	
OS	Zea mays.
XX	
FN	WO200009706-A2.
XX	
PD	24-FEB-2000.
XX	
PP	16-AUG-1999; 99WO-US18760.
XX	
PR	17-AUG-1998; 98US-0096822.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Dhuga KS, Helentjaris TG, Bowen BA, Wang X;
XX	
DR	WPI; 2000-224343/19.
XX	
NR	N-PSDB; AA299503.
XX	
PT	New genes which encode maize cellulose synthase polypeptides in plants
PT	useful for modulating the expression of cellulose synthase in plants

PT and to produce transgenic plants expressing the novel protein -  
 XX Claim 15; Page 118-120; 119pp; English.  
 XX The present sequence represents a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.  
 CC  
 SQ Sequence 1086 AA;  
 Query Match 100.0%; Score 5918; DB 21; Length 1086;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEASAGLVAGSHNRNELVYIRRDGPDPKPPREQNGVQICGDVGLAPGDPPVACNE 60  
 DB 1 MEASAGLVAGSHNRNELVYIRRDGPDPKPPREQNGVQICGDVGLAPGDPPVACNE 60  
 QY 61 CAPPCVRCDCYEYERREBGTQNCPOCKTRYRLKCCQAVTDEEDBDGVDLDNEPNMDGHS 120  
 DB 61 CAPPCVRCDCYEYERREBGTQNCPOCKTRYRLKCCQAVTDEEDBDGVDLDNEPNMDGHS 120  
 QY 121 GSVAESMLYGHMSYGGGDPNGAPQAFQALNPVPLTLTNGQWDDLPPEQHALVPSFMGG 180  
 DB 121 GSVAESMLYGHMSYGGGDPNGAPQAFQALNPVPLTLTNGQWDDLPPEQHALVPSFMGG 180  
 QY 181 GKRIHPLPYADPSLPVQPSMDPSKDLAAVGYGSVAMKERMENWKQQRMRHQTGNDGG 240  
 DB 181 GKRIHPLPYADPSLPVQPSMDPSKDLAAVGYGSVAMKERMENWKQQRMRHQTGNDGG 240  
 QY 241 DGGDDADLPMBARQQLSRKIPLPSSQINPYRMIIILVVLGFFTHYVMPVNDAPA 300  
 DB 241 DGGDDADLPMBARQQLSRKIPLPSSQINPYRMIIILVVLGFFTHYVMPVNDAPA 300  
 QY 301 LMLISVIGEIFWPMMSWILDOFPKWPPIERETVLDRLSLRFDKGGQSLAPIDFVSTVD 360  
 DB 301 LMLISVIGEIFWPMMSWILDOFPKWPPIERETVLDRLSLRFDKGGQSLAPIDFVSTVD 360  
 QY 361 PLKEPPLVTNTVLSLISVDYVDKVCVSDGAAMLTFEALSETSEPAKXWPFCKXY 420  
 DB 361 PLKEPPLVTNTVLSLISVDYVDKVCVSDGAAMLTFEALSETSEPAKXWPFCKXY 420  
 QY 421 NIEPRAPENVYFOOKIDYLDKQVAANFVRERAPMKREYEFKVRINALVAKQVPEEGT 480  
 DB 421 NIEPRAPENVYFOOKIDYLDKQVAANFVRERAPMKREYEFKVRINALVAKQVPEEGT 480  
 QY 481 MODGTFWPGNNVVDHFGMTQVFLGSGGGLDCGNEPLRLVYVSREKRPQYNNHKKAGANN 540  
 DB 481 MODGTFWPGNNVVDHFGMTQVFLGSGGGLDCGNEPLRLVYVSREKRPQYNNHKKAGANN 540  
 QY 541 ALVRSVAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGDIDR 600  
 DB 541 ALVRSVAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGDIDR 600  
 QY 601 HDRYANRNVFPDINMKGLDGIQGPITYGTGCVFRQALYGDAPKTKPSPRTCCMPK 660

DB 601 HDRYANRNVFPDINMKGLDGIQGPITYGTGCVFRQALYGDAPKTKPSPRTCCMPK 660  
 QY 661 WPCFCCCGNRRKOKKTKPKTEKKKLLFPKKEENQSPAYALGEIDEAPGAENEKAGIVN 720  
 DB 661 WPCFCCCGNRRKOKKTKPKTEKKKLLFPKKEENQSPAYALGEIDEAPGAENEKAGIVN 720  
 QY 721 QOKLEKRGQSSVPTSTLLENGTLKASAPASLILKEAIVHISGVEDKTDWKEIGWY 780  
 DB 721 QOKLEKRGQSSVPTSTLLENGTLKASAPASLILKEAIVHISGVEDKTDWKEIGWY 780  
 QY 781 GSVEDILITGKMHCHGRSITCYI PKRYPAGSA PLNLSDBLHGVWALGSIIFPSNH 840  
 DB 781 GSVEDILITGKMHCHGRSITCYI PKRYPAGSA PLNLSDBLHGVWALGSIIFPSNH 840  
 QY 841 CPLWYGGGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLTGKRTPELNNVASLW 900  
 DB 841 CPLWYGGGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLTGKRTPELNNVASLW 900  
 QY 901 FMSLFTICFATSIEMRWSGVGIDDMWRNEQFWIIGVSSHLPFAVQGLKVIAGVDTSP 960  
 DB 901 FMSLFTICFATSIEMRWSGVGIDDMWRNEQFWIIGVSSHLPFAVQGLKVIAGVDTSP 960  
 QY 961 TTSKGDDESESELYTEKWTLLIPPTLLILNPIGVAVSNAINNGVSWGPLEFGL 1020  
 DB 961 TTSKGDDESESELYTEKWTLLIPPTLLILNPIGVAVSNAINNGVSWGPLEFGL 1020  
 QY 1021 FFAFWVIVHLTPFLKGLVGRNRTPTIIVMSILASIFSLMWRIDPFLAKDGPFLLE 1080  
 DB 1021 FFAFWVIVHLTPFLKGLVGRNRTPTIIVMSILASIFSLMWRIDPFLAKDGPFLLE 1080  
 QY 1081 CGLDGN 1086  
 DB 1081 CGLDGN 1086  
 RESULT 2  
 AAY84116  
 ID AAY84116 standard; Protein; 1086 AA.  
 XX  
 AC AAY84116;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE Amino acid sequence of a maize cellulose synthase.  
 XX  
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
 XX transgenic plant; plant breeding marker.  
 OS  
 XX  
 PN Zea mays.  
 XX  
 WO200009706-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 16-AUG-1999; 99WO-US18760.  
 XX  
 PR 17-AUG-1998; 98US-0096822.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Dhuga KS, Helentjaris TG, Bowen BA, Wang X;  
 XX  
 DR WPI; 2000-224343/19.  
 XX  
 DR N-PSDB; AA299918.  
 PT  
 PT New genes which encode maize cellulose synthase polypeptides in plants  
 PT useful for modulating the expression of cellulose synthase in plants  
 PT and to produce transgenic plants expressing the novel protein -  
 PS Claim 15; Page 157-160; 119pp; English.  
 CC The present sequence represents a maize cellulose synthase polypeptide.

The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plants, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Sequence 1086 AA;

Query Match 100.0%; Score 5918; DB 21; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEASAGLVAGSHNRELVIIRDDGPGKPPBQNGVCOICGDDVGLAPGDDPFAVNE 60  
1 MEASAGLVAGSHNRELVIIRDDGPGKPPBQNGVCOICGDDVGLAPGDDPFAVNE 60  
CAPFPCRCYEREREGTQNGPQCKTRKYLKGCQRTGDEEDVDLDNEFMNDGHS 120  
CAPFPCRCYEREREGTQNGPQCKTRKYLKGCQRTGDEEDVDLDNEFMNDGHS 120  
QSVASMLYGHMSYGRGDPNGAPAPQAPQINPVPLLTNGQWVDIPEQHALVPSFMGGG 180  
QSVASMLYGHMSYGRGDPNGAPAPQAPQINPVPLLTNGQWVDIPEQHALVPSFMGGG 180  
QSVASMLYGHMSYGRGDPNGAPAPQAPQINPVPLLTNGQWVDIPEQHALVPSFMGGG 180  
QSVASMLYGHMSYGRGDPNGAPAPQAPQINPVPLLTNGQWVDIPEQHALVPSFMGGG 180  
GKRTHPLPYADPSLEVPQPSMDPSKDLAAYGYGVAMKERMENMKQGERMHTQNDGGG 240  
GKRTHPLPYADPSLEVPQPSMDPSKDLAAYGYGVAMKERMENMKQGERMHTQNDGGG 240  
GKRTHPLPYADPSLEVPQPSMDPSKDLAAYGYGVAMKERMENMKQGERMHTQNDGGG 240  
GKRTHPLPYADPSLEVPQPSMDPSKDLAAYGYGVAMKERMENMKQGERMHTQNDGGG 240  
DDGDDADPLMDEARQSLRKILPSSQINPYRMIIIRLVYGFYFRYRMHPNDFA 300  
DDGDDADPLMDEARQSLRKILPSSQINPYRMIIIRLVYGFYFRYRMHPNDFA 300  
DDGDDADPLMDEARQSLRKILPSSQINPYRMIIIRLVYGFYFRYRMHPNDFA 300  
DDGDDADPLMDEARQSLRKILPSSQINPYRMIIIRLVYGFYFRYRMHPNDFA 300  
LMLISVCEIWFAMSWIIDQFPKPFIERETYLDRLSRFDKEGQPSGLAIFDFFVSTVD 360  
LMLISVCEIWFAMSWIIDQFPKPFIERETYLDRLSRFDKEGQPSGLAIFDFFVSTVD 360  
LMLISVCEIWFAMSWIIDQFPKPFIERETYLDRLSRFDKEGQPSGLAIFDFFVSTVD 360  
LMLISVCEIWFAMSWIIDQFPKPFIERETYLDRLSRFDKEGQPSGLAIFDFFVSTVD 360  
PLKEPPLVTNTVLSISVDYPVDKSCVSDGAAMLTFEALSTSEFAKMWPFCKRY 420  
PLKEPPLVTNTVLSISVDYPVDKSCVSDGAAMLTFEALSTSEFAKMWPFCKRY 420  
PLKEPPLVTNTVLSISVDYPVDKSCVSDGAAMLTFEALSTSEFAKMWPFCKRY 420  
PLKEPPLVTNTVLSISVDYPVDKSCVSDGAAMLTFEALSTSEFAKMWPFCKRY 420  
MODGTPMPGNVVRBHPGMIQVFLGSGGDLCEGNEPLPLVYVSRKRGYNNHKKAGAMN 540  
MODGTPMPGNVVRBHPGMIQVFLGSGGDLCEGNEPLPLVYVSRKRGYNNHKKAGAMN 540  
MODGTPMPGNVVRBHPGMIQVFLGSGGDLCEGNEPLPLVYVSRKRGYNNHKKAGAMN 540  
MODGTPMPGNVVRBHPGMIQVFLGSGGDLCEGNEPLPLVYVSRKRGYNNHKKAGAMN 540  
ALVRSAVLTNAPYLLNDCDHYINNSKAKEAMCFMMDPLLGKKVCVVGQFQRDGDIDR 600  
ALVRSAVLTNAPYLLNDCDHYINNSKAKEAMCFMMDPLLGKKVCVVGQFQRDGDIDR 600  
ALVRSAVLTNAPYLLNDCDHYINNSKAKEAMCFMMDPLLGKKVCVVGQFQRDGDIDR 600  
ALVRSAVLTNAPYLLNDCDHYINNSKAKEAMCFMMDPLLGKKVCVVGQFQRDGDIDR 600  
HDRVANRVNVPDIDMKGLDGIQGPITYGTGCVFRRQALLGYDAKTKKPSRTNCMPK 660  
HDRVANRVNVPDIDMKGLDGIQGPITYGTGCVFRRQALLGYDAKTKKPSRTNCMPK 660  
HDRVANRVNVPDIDMKGLDGIQGPITYGTGCVFRRQALLGYDAKTKKPSRTNCMPK 660  
HDRVANRVNVPDIDMKGLDGIQGPITYGTGCVFRRQALLGYDAKTKKPSRTNCMPK 660  
WCFCGCCGGRKOKTKTKTEKKKLLFFKKKEBNSPYVALGEIDEAPGAENEKAGIVN 720  
WCFCGCCGGRKOKTKTKTEKKKLLFFKKKEBNSPYVALGEIDEAPGAENEKAGIVN 720  
WCFCGCCGGRKOKTKTKTEKKKLLFFKKKEBNSPYVALGEIDEAPGAENEKAGIVN 720  
WCFCGCCGGRKOKTKTKTEKKKLLFFKKKEBNSPYVALGEIDEAPGAENEKAGIVN 720

661 WCFCGCCGGRKOKTKTKTEKKKLLFFKKKEBNSPYVALGEIDEAPGAENEKAGIVN 720  
QOKLEKFGGSSVVTSTLENQGLKASGASLKLKALHIVISGCEYEDKTDWKEIGWY 780  
QOKLEKFGGSSVVTSTLENQGLKASGASLKLKALHIVISGCEYEDKTDWKEIGWY 780  
QOKLEKFGGSSVVTSTLENQGLKASGASLKLKALHIVISGCEYEDKTDWKEIGWY 780  
QOKLEKFGGSSVVTSTLENQGLKASGASLKLKALHIVISGCEYEDKTDWKEIGWY 780  
GSVTEEDILTGKQNHCHMRSTYCPKRVAFGSAPLNLSRHLQVLMWALGSIIFPSNH 840  
GSVTEEDILTGKQNHCHMRSTYCPKRVAFGSAPLNLSRHLQVLMWALGSIIFPSNH 840  
GSVTEEDILTGKQNHCHMRSTYCPKRVAFGSAPLNLSRHLQVLMWALGSIIFPSNH 840  
GSVTEEDILTGKQNHCHMRSTYCPKRVAFGSAPLNLSRHLQVLMWALGSIIFPSNH 840  
CPLWYGVGGIKFLERSYSINSIYVPTSTIPLAAYCTLPALICLTGKFTPELNNVASLW 900  
CPLWYGVGGIKFLERSYSINSIYVPTSTIPLAAYCTLPALICLTGKFTPELNNVASLW 900  
CPLWYGVGGIKFLERSYSINSIYVPTSTIPLAAYCTLPALICLTGKFTPELNNVASLW 900  
CPLWYGVGGIKFLERSYSINSIYVPTSTIPLAAYCTLPALICLTGKFTPELNNVASLW 900  
FMSLFIPIFATSIEMMSGVDIDMWRNEQFWIIGVSSHSLFAVQGLLVAGVDTSF 960  
FMSLFIPIFATSIEMMSGVDIDMWRNEQFWIIGVSSHSLFAVQGLLVAGVDTSF 960  
FMSLFIPIFATSIEMMSGVDIDMWRNEQFWIIGVSSHSLFAVQGLLVAGVDTSF 960  
FMSLFIPIFATSIEMMSGVDIDMWRNEQFWIIGVSSHSLFAVQGLLVAGVDTSF 960  
TVTSKGGDDEFSRLYFKMTTLLIPPTTLLNFIQVAVGSAINNYESMGPFLGKL 1020  
TVTSKGGDDEFSRLYFKMTTLLIPPTTLLNFIQVAVGSAINNYESMGPFLGKL 1020  
TVTSKGGDDEFSRLYFKMTTLLIPPTTLLNFIQVAVGSAINNYESMGPFLGKL 1020  
TVTSKGGDDEFSRLYFKMTTLLIPPTTLLNFIQVAVGSAINNYESMGPFLGKL 1020  
FFAFWIVHLVYPLKGLVGRQNRPTIVIVMSILLASIFSLWVRIDPFLAKDGPFLER 1080  
FFAFWIVHLVYPLKGLVGRQNRPTIVIVMSILLASIFSLWVRIDPFLAKDGPFLER 1080  
FFAFWIVHLVYPLKGLVGRQNRPTIVIVMSILLASIFSLWVRIDPFLAKDGPFLER 1080  
FFAFWIVHLVYPLKGLVGRQNRPTIVIVMSILLASIFSLWVRIDPFLAKDGPFLER 1080  
CGLDGN 1086  
CGLDGN 1086  
CGLDGN 1086  
CGLDGN 1086

RESULT 3  
AAY58835  
ID AAY58835 standard; Protein: 1086 AA.  
AAY58835;  
AAY58835;  
08-MAY-2000 (first entry)  
Corn cellulose synthase (full-length).  
Corn; maize; cellulose synthase; transgenic plant.  
Zea mays.  
W0200004166-A2.  
27-JAN-2000.  
13-JUL-1999; 99WO-US15871.  
14-JUL-1998; 98US-0092844.  
(DUPO) DU PONT DE NEMOURS & CO E I.  
Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G, Ratalaki JA, Thorpe CJ;  
MPI: 2000-182431/16.  
N-PSDB; AAZ58266.  
Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as probes for isolating cDNAs and genes encoding homologous proteins, for producing transgenic plants -  
Claim 12; Page 50-53; 93pp; English.  
The present sequence is that of full-length corn cellulose synthase (CS), as deduced from an isolated cDNA clone (see AAY58266). The invention relates to isolated CS nucleic acid fragments and to CS polypeptides. It also relates to the construction of a chimeric gene encoding all or a portion of the CS, in sense or antisense orientation, where expression of the gene results in altered levels

of the CS in transformed host cells. The host cells can be used to screen compounds for their ability to inhibit CS activity. CS nucleic acids are also useful for producing transgenic plants having altered levels of CS, and hence altered levels of fibre. CS may also serve as a target for the development of novel herbicides.

XX Sequence 1086 AA;

Query Match 100.0%; Score 5918; DB 21; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEASAGLVAGSHNRNELVIRRDGPGRPREQNGVCQICGGDVGLAPGGDPVACNE 60
DB 1 MEASAGLVAGSHNRNELVIRRDGPGRPREQNGVCQICGGDVGLAPGGDPVACNE 60
QY 61 CAPVCRDCYERERREGTONCPOCKTRYRLKGCQKVTGDEBEDDGDLDNEFNMDGHDS 120
DB 61 CAPVCRDCYERERREGTONCPOCKTRYRLKGCQKVTGDEBEDDGDLDNEFNMDGHDS 120
QY 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLITNGQMYDDIPEEQHALVPSFMGG 180
DB 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLITNGQMYDDIPEEQHALVPSFMGG 180
QY 181 GRIRHPLPYADPSLPVQPRSMDSKDLAAYGGSVAMKEREMWKORQRMHQGTGDCGG 240
DB 181 GRIRHPLPYADPSLPVQPRSMDSKDLAAYGGSVAMKEREMWKORQRMHQGTGDCGG 240
QY 241 DGGDDADLPLMEARQOLSRKIPLPSQINPNRMIIIFLVLGPFPHRVHNPVADAF 300
DB 241 DGGDDADLPLMEARQOLSRKIPLPSQINPNRMIIIFLVLGPFPHRVHNPVADAF 300
QY 301 LMLISYICBIFWPMWILDOFPKPFIERETYLRLSLRFDKSGQSPQLAIDFVSTVD 360
DB 301 LMLISYICBIFWPMWILDOFPKPFIERETYLRLSLRFDKSGQSPQLAIDFVSTVD 360
QY 361 PLKEPPLVTNTVLSLSDVYPVDKVSQVSDGAAMLTPEALSETSEPAKKNVPCRY 420
DB 361 PLKEPPLVTNTVLSLSDVYPVDKVSQVSDGAAMLTPEALSETSEPAKKNVPCRY 420
QY 421 NLEPRAPBMYFOCKIDYLDKDVANFVRBRAMKREYEEFKYRINLVAKQVPEEGT 480
DB 421 NLEPRAPBMYFOCKIDYLDKDVANFVRBRAMKREYEEFKYRINLVAKQVPEEGT 480
QY 481 MODGTMPGNVNDHFGMIQVFLGSGGLDCEGNELEPRLVYYSREKRPQGNHKKAGAN 540
DB 481 MODGTMPGNVNDHFGMIQVFLGSGGLDCEGNELEPRLVYYSREKRPQGNHKKAGAN 540
QY 541 ALVRSASVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGD 600
DB 541 ALVRSASVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGD 600
QY 601 HGRVARNVFPINMKGLDGOPIYVNGCVFRQALYGVAPRTKXKPRSTQCMWR 660
DB 601 HGRVARNVFPINMKGLDGOPIYVNGCVFRQALYGVAPRTKXKPRSTQCMWR 660
QY 661 WCFCCCFGNRKOKTKTKTEKKKLLFKKEENQSPAYALGEIDEAAGABENKAGIYN 720
DB 661 WCFCCCFGNRKOKTKTKTEKKKLLFKKEENQSPAYALGEIDEAAGABENKAGIYN 720
QY 721 QOKLEKKFGQSSVFTSTLLENGGTLKSASPASLKEAIVHISCGYEDKTDWKEIGMI 780
DB 721 QOKLEKKFGQSSVFTSTLLENGGTLKSASPASLKEAIVHISCGYEDKTDWKEIGMI 780
QY 781 GSVTEBILGFKKHCGRMSIYCIIPKRVAFKGSAPNLSPRLHOUVPMALGSEIFFSNH 840
DB 781 GSVTEBILGFKKHCGRMSIYCIIPKRVAFKGSAPNLSPRLHOUVPMALGSEIFFSNH 840
QY 841 CPLMYGGGLKLEKLEFSYINSIYVPMTSIPLLAYCTLPACILTKGKFTPELANTASIM 900
DB 841 CPLMYGGGLKLEKLEFSYINSIYVPMTSIPLLAYCTLPACILTKGKFTPELANTASIM 900
QY 901 FMSLFICIFATSTILEKRWGSGVIDDWMRNEQFVAVIGVSSHLPVAFQGLLKVIAGVDTSF 960

```

```

DB 901 FMSLFICIFATSTILEKRWGSGVIDDWMRNEQFVAVIGVSSHLPVAFQGLLKVIAGVDTSF 960
QY 961 TVTSKGGDEEFSELYFKMTTLLIPTTLLNFIQVAVSNAINGYESMPLFGKL 1020
DB 961 TVTSKGGDEEFSELYFKMTTLLIPTTLLNFIQVAVSNAINGYESMPLFGKL 1020
QY 1021 FFAFWIVHLVPLKGLVGRNRTPTIYVMSIILASIFSLMWRIDPFLAKDGPILIE 1080
DB 1021 FFAFWIVHLVPLKGLVGRNRTPTIYVMSIILASIFSLMWRIDPFLAKDGPILIE 1080
QY 1081 GGLDCN 1086
DB 1081 GGLDCN 1086

```

```

RESULT 4
ID AAY58832 standard; Protein, 1148 AA.
AC AAY58832;
XX 08-MAY-2000 (first entry)
DT Corn cellulose synthase (partial sequence).
DE Corn; maize; cellulose synthase; transgenic plant.
XX Zea mays.
OS WO200004166-A2.
PD 27-JAN-2000.
PF 13-JUL-1999; 99WO-US15871.
PR 14-JUL-1998; 98US-0092844.
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G,
PI Rafaelski JA, Thorpe CJ;
DR WPI; 2000-182431/16.
DR N-PSDB; AA258263.
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
PS Claim 12; Page 39-42; 93pp; English.
XX The present sequence is that of a portion of corn cellulose
XX synthase (CS), as deduced from a contig of cDNA clones (see AAY58263).
XX The invention relates to isolated CS nucleic acid fragments and to
XX CS polypeptides. It also relates to the construction of a chimeric
XX gene encoding all or a portion of the CS, in sense or antisense
XX orientation, where expression of the gene results in altered levels
XX of the CS in transformed host cells. The host cells can be used to
XX screen compounds for their ability to inhibit CS activity. CS
XX nucleic acids are also useful for producing transgenic plants having
XX altered levels of CS, and hence altered levels of fibre. CS may
XX also serve as a target for the development of novel herbicides.
SQ Sequence 1148 AA;
Query Match 94.1%; Score 5570.5; DB 21; Length 1148;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1027; Conservative 23; Mismatches 35; Indels 5; Gaps 3;
QY 1 MEASAGLVAGSHNRNELVIRRDGPGRPREQNGVCQICGGDVGLAPGGDPVACNE 60
DB 60 MEASAGLVAGSHNRNELVIRRDGPGRPREQNGVCQICGGDVGLAPGGDPVACNE 119

```





```

QY 152 NVPLLTNGOMVDIPPEQHALVPSFMGGGKRIHPYADPSLPVQPSGMDPSKDLAAVG 211
DB 123 NVPLLTNGOMVDIPPEQHALVPSFMGGGKRIHPYADPSLPVQPSGMDPSKDLAAVG 182
QY 212 YGSVAMKERNEMKROERHQTGNDGGDDGADALPLMDEARQOLSRKIPSPSQINP 271
DB 183 YGSVAMKERNEMKROERHQTGNDGGDDGADALPLMDEARQOLSRKIPSPSQINP 242
QY 272 YRMIIIRLVIGFPFHYRVMHPVNDALFALMLISVICIWMFAMSWILDGPKMFIPIRET 311
DB 243 YRMIIIRLVIGFPFHYRVMHPVNDALFALMLISVICIWMFAMSWILDGPKMFIPIRET 302
QY 332 YLDRLRLRPDKEQPSQALPIDEFVSTVDPLEKPPVLTNTVTSIIISVDYVDKSCYYS 391
DB 303 YLDRLRLRPDKEQPSQALPIDEFVSTVDPLEKPPVLTNTVTSIIISVDYVDKSCYYS 362
QY 392 DQGAAMLTPEALSETSPFAKKWVPCKRYNIEPRAPEMWYQOKIDYLKQDVANFRRER 451
DB 363 DQGAAMLTPEALSETSPFAKKWVPCKRYNIEPRAPEMWYQOKIDYLKQDVANFRRER 422
QY 452 AMREYEEFKVRINALVAKAKVPEBGMTWQDGPWGNVNDHPGMIQVFLGSGGLDC 511
DB 423 AMREYEEFKVRINALVAKAKVPEBGMTWQDGPWGNVNDHPGMIQVFLGSGGLDC 482
QY 512 EGNELPRLVYVSREKPGYNNHKKAGAMNALVRSVAVLTAPVTLNDCDHYINNSKATK 571
DB 483 EGNELPRLVYVSREKPGYNNHKKAGAMNALVRSVAVLTAPVTLNDCDHYINNSKATK 542
QY 572 EAMCFMMDPLGKKVVCVOPFORPDGIDRDRYANRVVFPDINMGKLOGIPVYVG 631
DB 543 EAMCFMMDPLGKKVVCVOPFORPDGIDRDRYANRVVFPDINMGKLOGIPVYVG 602
QY 632 CVRRQALYGYDAPKTKKPPSRTCNCPKWCFCGCCGNNKOK---KTKPKTEKKKLLF 688
DB 603 CVRRQALYGYDAPKTKKPPSRTCNCPKWCFCGCCGNNKOK---KTKPKTEKKKLLF 661
QY 689 FKKEENOSPAYALGEIDEAPGABENKAGIVNOOKLEKKGSSVFTSTLLENGTLLS 748
DB 662 FKKEENOSPAYALGEIDEAPGABENKAGIVNOOKLEKKGSSVFTSTLLENGTLLS 721
QY 749 ASPASILLKEAIIHISGCEYEDKTOMGKEIGWYGSVTEIDITGPKMCHGRSITCIPIKRV 808
DB 722 ASPASILLKEAIIHISGCEYEDKTOMGKEIGWYGSVTEIDITGPKMCHGRSITCIPIKRV 781
QY 809 AFGSAPLNLSDRLHQVLRWALGSIIEIFSNHCPLMWYGGGLKFLERSYINSIYVPT 868
DB 782 AFGSAPLNLSDRLHQVLRWALGSIIEIFSNHCPLMWYGGGLKFLERSYINSIYVPT 841
QY 869 SIPLLAYCTLPALICLTGKFTTPELNNVASLWMSLFCIFATSIEMRSGVIGIDWNR 928
DB 842 SIPLLAYCTLPALICLTGKFTTPELNNVASLWMSLFCIFATSIEMRSGVIGIDWNR 901
QY 929 NEOPWYIGVSHLPAVFOGLLVYIAGVDSFTWTSISGSDSESELYTKMTTLIIPPT 988
DB 902 NEOPWYIGVSHLPAVFOGLLVYIAGVDSFTWTSISGSDSESELYTKMTTLIIPPT 961
QY 989 TLLLLNFIQVAVASNAINNYESWGPLFGKLFPAFVNIYHLYPFLKGLVGRQNRPTTV 1048
DB 962 TLLLLNFIQVAVASNAINNYESWGPLFGKLFPAFVNIYHLYPFLKGLVGRQNRPTTV 1021
QY 1049 IVMSILLASIFSLWVRIDPPLAKDGPILLECGGLDC 1085
DB 1022 IVMSILLASIFSLWVRIDPPLAKDGPILLECGGLDC 1058

```

RESULT 6  
ID AAY84113 standard; Protein; 1094 AA.  
XX AAY84113;  
XX  
DT 03-JUL-2000 (first entry)  
XX

```

DE Amino acid sequence of a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
OS Zea mays.
PN WO200009706-A2.
PD 24-FEB-2000.
PF 16-AUG-1999; 99W0-US18760.
PK 17-AUG-1998; 98US-0096822.
PR (PION-) PIONEER HI-BRED INT INC.
PA Dhuga KS, Helentjaris TG, Bowen BA, Wang X;
PI WPI, 2000-224343/19.
PX N-PSDB; AA29509.
DR
XX
XX The present sequence represents a maize cellulose synthase polypeptide.
XX The cellulose synthase can be used for the improvement of stalk quality
XX for improved stand or silage. It also provides an increased concentration
XX of cellulose in the pericarp, hardening the kernel and improving the
XX handling ability. The sequences are used to produce transgenic plants
XX and seeds expressing the cellulose synthase. The polynucleotide is
XX used for modulating, preferably increasing, the level of the synthase
XX in a plant cell. The plants are preferably monocots. The polynucleotide
XX is also used as a probe or primer in the detection quantitation or
XX isolation of gene transcripts. The probes are useful in detecting
XX deficiencies in the level of mRNA in screenings for desired transgenic
XX plant, for detecting mutations in the gene, for monitoring upregulation
XX of expression or changes in enzyme activity in screening assays of
XX compounds, for detection of any number of allelic variants of the gene,
XX or for use as molecular markers in plant breeding programs. The
XX isolated nucleic acids of the present invention can also be used for
XX recombinant expression of their encoded polypeptides or for use as
XX immunogens in the preparation and/or screening of antibodies. The
XX proteins can be employed in assays for enzyme agonists or antagonists
XX of enzyme function or for use of immunogens or antigens to obtain
XX antibodies specifically immunoreactive with a protein.
SQ Sequence 1094 AA;
Query Match 88.3%; Score 5228; DB 21; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;
QY 1 MEASAGLVAGSHNRRELVVIRRD---GDGPKPPEQNGVQICGDDVGLAGGDPFVA 57
DB 1 MEASAGLVAGSHNRRELVVIRDRRESGAAGGAARAAEP-CQIGDEVGVFDESPFA 59
QY 58 CNECAFPVRCDCYERBRGTONCQCKTRYRLKGCQAVTDBEDDEYDIDLDFN-ND 116
DB 60 CNECAFPVRCACYERBRGSGQACPCQCRTRYRLKGCPCPVADBEDDEYDIDLEGGGLD 119
QY 117 G----HDSQVASEMLYGHMSYGRGDPNGAPQAFOLANNVPLLTNGQVNDIPPEQHAL 172
DB 120 GAHEDDPQYVAESMLRAQMSYGRGD---AHPGSPVNVNVLNTNGQVNDIPPEQHAL 176
QY 173 VPSFM---GGGKRIHPLYADPSLPVQPSRMDPSKDLAAVYGSVAMKERNEMKROER 229
DB 177 VPSYMGGGGKRIHPLEFADPNLPVQPSRMDPSKDLAAVYGSVAMKERNEMKROER 236
QY 230 RMHQTONDGGD-DGDDALPLMDEARQOLSRKIPSPSQINPYRMIIIRLVIGFPH 288

```

```

Db      227  RLQHYRSBSGGDWDDADLPLMDARQLSRKVPISRRINPMMIIVIRLVVGLGFFPH 296
Qy      289  YRVHAPVNDAPALMLISVCEIWMFAMSWILDOFPKMFPIERTYLDRLSLRDEKQSPSQ 348
Db      297  YRVHAPADAFALMLISVCEIWMFAMSWILDOFPKMLPIERTYLDRLSLRDEKQSPSQ 356
Qy      349  LAPIDFVSTVDPLKEPPLVTTNTVLSISVDYPRVDKVSQYSDGAMLTPEALSETSE 408
Db      357  LAPIDFVSTVDPLKEPPLVTTNTVLSISVDYPRVDKVSQYSDGAMLTPEALSETSE 416
Qy      409  FAKKVVPPCKRYNIPRAPEWYFOQKIDYLDKQVAAFREBRAMKREEEKVAINLV 468
Db      417  FAKKVVPPCKRYNIPRAPEWYFOQKIDYLDKQVAAFREBRAMKREEEKVAINLV 476
Qy      469  AKAQKVPBEGMTMODGTPWPGANNVBDHPGMIQVFLQSGSGGLDCEGNEJLPRLYVYREKRP 528
Db      477  AKAQKVPBEGMTMODGTPWPGANNVBDHPGMIQVFLQSGSGGRVNEGNEJLPRLYVYREKRP 536
Qy      539  GYNHKKAGAMALVRSAYLTNAPYLLNDCDHYIINNSKAIKEAMCFMMDPLGKRCY 588
Db      537  GYNHKKAGAMALVRSAYLTNAPYLLNDCDHYIINNSKAIKEAMCFMMDPLGKRCY 596
Qy      569  VQFPORFOSIDRHDYANRNVPFIDIMKGLDGIQGIPIVGTGCVFRQALGYDAPKTK 648
Db      597  VQFPORFOSIDRHDYANRNVPFIDIMKGLDGIQGIPIVGTGCVFRQALGYDAPKTK 656
Qy      649  KPPSRNCNWPKMCFCFCCCFGNRKQKTKPKTEKKKLFFKKEENSPAYALGEIDEA 708
Db      657  KPPSRNCNWPKMCFCFCCCFGNRKQKTKPKTEKKKLFFKKEENSPAYALGEIDEA 716
Qy      709  PGAENKAGIVNQOKLEKKGQSVFVSTLLENGTILKASAPASILKEAIIHVISGYED 768
Db      717  PGADIEKAGIVNQOKLEKKGQSVFVSTLLENGTILKASAPASILKEAIIHVISGYED 776
Qy      769  KTDWKEIGMITYGSTEDILTFPKHCHWRSTIYCIPIKVAKGSAPLNSRLQVLRM 828
Db      777  KTDWKEIGMITYGSTEDILTFPKHCHWRSTIYCIPIKVAKGSAPLNSRLQVLRM 836
Qy      829  ALGSIIEFPNSHCPLWYGGGLKFLERSYINSIYPMWTSIPLAYCTLPATCLLTGKF 888
Db      837  ALGSIIEFPNSHCPLWYGGGLKFLERSYINSIYPMWTSIPLAYCTLPATCLLTGKF 896
Qy      889  ITPBLNNVASLWMSLFICIFATSILEMMSGVGIDDMWRNDFWVIGVSSHLPVAFQ 948
Db      897  ITPBLNNVASLWMSLFICIFATSILEMMSGVGIDDMWRNDFWVIGVSSHLPVAFQ 956
Qy      949  LUKVAGVDTSTFTVTSKGGDEEFSELYTFKWTLLIPTTLLLNFTIVVAGVSNAINN 1008
Db      957  LUKVAGVDTSTFTVTSKGGDEEFSELYTFKWTLLIPTTLLLNFTIVVAGVSNAINN 1016
Qy      1009  GYESMGPLRGLFPAFWIVHLYPFLKGLVGRONTPTIVVMSILLASIFSLWVRIDP 1068
Db      1017  GYESMGPLRGLFPAFWIVHLYPFLKGLVGRONTPTIVVMSILLASIFSLWVRIDP 1076
Qy      1069  FLAKODGPLLEEGGLDCN 1086
Db      1077  FLAKSNGPLLEEGGLDCN 1094

```

```

RESULT 7
AAy84118
ID      AAy84118 standard; Protein; 1094 AA.
XX      AAy84118;
XX      03-JUL-2000 (first entry)
XX      Amino acid sequence of a maize cellulose synthase.
XX      Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
XX      transgenic plant; plant breeding marker.
XX

```

```

OS      Zea mays.
FN      WO200009706-A2.
XX      24-FEB-2000.
XX      16-AUG-1999; 99WO-US18760.
XX      17-AUG-1998; 98US-0096822.
XX      (PION-) PIONEER HI-BRED INT INC.
XX      Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX      WPI; 2000-224343/19.
XX      N-PSDB; AA299524.
XX      New genes which encode maize cellulose synthase polypeptides in plants
XX      useful for modulating the expression of cellulose synthase in plants
XX      and to produce transgenic plants expressing the novel protein
XX      Claim 15; Page 173-175; 119pp; English.
XX
XX      The present sequence represents a maize cellulose synthase polypeptide.
XX      The cellulose synthase can be used for the improvement of stalk quality
XX      for improved stand or silage. It also provides an increased concentration
XX      of cellulose in the pericarp, hardening the kernel and improving its
XX      handling ability. The sequences are used to produce transgenic plants
XX      and seeds expressing the cellulose synthase. The polynucleotide is
XX      used for modulating, preferably increasing, the level of the synthase
XX      in a plant cell. The plants are preferably monocots. The polynucleotide
XX      is also used as a probe or primer in the detection quantitation or
XX      isolation of gene transcripts. The probes are useful in detecting
XX      deficiencies in the level of mRNA in screenings for desired transgenic
XX      plant, for detecting mutations in the gene, for monitoring upregulation
XX      of expression or changes in enzyme activity in screening assays of
XX      compounds, for detection of any number of allelic variants of the gene,
XX      or for use as molecular markers in plant breeding programs. The
XX      isolated nucleic acids of the present invention can also be used for
XX      recombinant expression of their encoded polypeptides or for use as
XX      immunogens in the preparation and/or screening of antibodies. The
XX      proteins can be employed in assays for enzyme agonists or antagonists
XX      of enzyme function or for use of immunogens or antigens to obtain
XX      antibodies specifically immunoreactive with a protein.
XX
XX      Sequence 1094 AA;
XX
XX      Query Match 88.3%; Score 5228; DB 21; Length 1094;
XX      Best Local Similarity 88.7%; Pred. No. 0;
XX      Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

```

```

Qy      1 MEASAGLVAGSHNRRLVIRRD--GDPGPKPREQNGQYQICIGDDVGLAPGDDPVA 57
Db      1 MEASAGLVAGSHNRRLVIRRDRESGAAGGAARRAP-CQICGDDVGLGFFDSFVA 59
Qy      58 CNECAFVYCRDCYERREBGTQNCPCQKTRYLKGCORVYGDDEEDVDLDLNEFN-WD 116
Db      60 CNECAFVYCRDCYERREBGTQNCPCQKTRYLKGCORVYGDDEEDVDLDLNEFN-WD 119
Qy      117 G---HDSQVAESMLYGHMSYGRGDPNGAPQAFQLNPNTPLLTNGQVDDIPRQIAL 172
Db      120 GAHEDDEPQYVAESMLRAQMSYGRGD--AHGFSPPVNPVPLLTNGQVDDIPRQIAL 176
Qy      173 VPSFW---GGGKRIHPIPVADPSLPVQPRGMDPEKDLAAGYGVAKEREMMKOROE 229
Db      177 VPSYSGGGGKRIHPIPVADPSLPVQPRGMDPEKDLAAGYGVAKEREMMKOROE 236
Qy      230 RMHQGTNDGGGD-DGDDADLPLMDARQOLSRKIPLPSSQINPYRMIIIRLVVGLGFFPH 288
Db      237 RLQHYRSBSGGDWDDADLPLMDARQPLSRKVPISRRINPMMIIVIRLVVGLGFFPH 296
Qy      289 YRVHAPVNDAPALMLISVCEIWMFAMSWILDOFPKMFPIERTYLDRLSLRDEKQSPSQ 348

```

Db 297 YRWHPADAFALMLISVICIWFAMSWILDOFPKMLPIERTTYLDRLSLRDKGQSQ 356  
 Qy 349 LAPIDFVSTVDPLKEPPLVTTNTVLSILSVDPYDKVSCVYSDGAAMLTFEALSETSE 408  
 Db 357 LAPIDFVSTVDPTKEPPLVTANTVLSILSVDPYDKVSCVYSDGAAMLTFEALSETSE 416  
 Qy 409 FAKKVPFCKRYNIEPRAPENYFOOKIDYLDKQVAAANFRRERAMKREYEERKYNALV 468  
 Db 417 FAKKVPFCKRYNIEPRAPENYFOOKIDYLDKQVAAANFRRERAMKREYEERKYNALV 476  
 Qy 469 AKAQVPEEGMTMOGTPMGPANNVBDHPGMIQVFLGQSGGRLCEGNEELRLVYVREKRP 528  
 Db 477 AKAQVPEEGMTMOGTPMGPANNVBDHPGMIQVFLGQSGGRLCEGNEELRLVYVREKRP 536  
 Qy 529 GYNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDLVGKVCY 588  
 Db 537 GYNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDLVGKVCY 596  
 Qy 589 VQPPQRPDGIHHRDRIANRVVFPDINMKGLDGIQGIYVGCYFRQALYGYDAPTK 648  
 Db 597 VQPPQRPDGIHHRDRIANRVVFPDINMKGLDGIQGIYVGCYFRQALYGYDAPTK 656  
 Qy 649 KRPSTCNQWPKMFCFCCCFGNRKOKTKTKPKTEKKXLLFFKKEENQSPAYALGEIDEA 708  
 Db 657 KRPSTCNQWPKMFCFCCCFGNRKOKTKTKPKTEKKXLLFFKKEENQSPAYALGEIDEA 716  
 Qy 709 PGAENEKAGIVNOQKLEKFGQSSVFTSTLENGTLLKASAPASILKEAIIHVISGVED 768  
 Db 717 PGADIEKAGIVNOQKLEKFGQSSVFTSTLENGTLLKASAPASILKEAIIHVISGVED 776  
 Qy 769 KTDWKEIGWYGSVTEEDILTGPKMCHGMSIYCI PKKVAFKGSAPLNLSRLHOVLRW 828  
 Db 777 KTDWKEIGWYGSVTEEDILTGPKMCHGMSIYCI PKKVAFKGSAPLNLSRLHOVLRW 836  
 Qy 829 ALGSTEIPEFSNHCPLWYGCGGLKLEERSYINSTIYYPWTSLPLAYCTLPALICLLTGKF 888  
 Db 837 ALGSTEIPEFSNHCPLWYGCGGLKLEERSYINSTIYYPWTSLPLAYCTLPALICLLTGKF 896  
 Qy 889 ITPELNNVASLWFMSLFICIFATSLIEMRSGVGLDMMWRNEQFWYIGVSSHLPVFCG 948  
 Db 897 ITPELNNVASLWFMSLFICIFATSLIEMRSGVGLDMMWRNEQFWYIGVSSHLPVFCG 956  
 Qy 949 LIAKTAGVDTSTVTSSKGGDEBFSELYFPKWTLLIIPPTLLILNFICVAVGSNAIIN 1008  
 Db 957 LIAKTAGVDTSTVTSSKGGDEBFSELYFPKWTLLIIPPTLLILNFICVAVGSNAIIN 1016  
 Qy 1009 GYESNGPLFEGKLFPFAFWIVLHYPRLKGLVGRONRPTTVIWSITLLASIFSLMWRIDP 1068  
 Db 1017 GYESNGPLFEGKLFPFAFWIVLHYPRLKGLVGRONRPTTVIWSITLLASIFSLMWRIDP 1076  
 Qy 1069 FLAKDGPILLECGGLDCN 1086  
 Db 1077 FLAKDGPILLECGGLDCN 1094  
 RESULT 8  
 AAY58834  
 ID AAY58834 standard; Protein; 1165 AA.  
 XX  
 AC AAY58834;  
 XX  
 DT 08-MAY-2000 (first entry)  
 XX  
 DE Corn cellulose synthase (partial sequence).  
 XX  
 KM Corn; maize; cellulose synthase; transgenic plant.  
 OS Zea mays.  
 PN WO200004166-A2.  
 XX  
 PD 27-JAN-2000.

PF 13-JUL-1999; 99MO-US15871.  
 XX  
 PR 14-JUL-1998; 98US-0092844.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.  
 XX  
 PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;  
 PI Rataleki JA, Thorpe CU;  
 XX  
 DR WPI; 2000-182431/16.  
 XX  
 DR N-PSDB; AA258265.  
 XX  
 PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
 PT probes for isolating cDNAs and genes encoding homologous proteins, for  
 PS producing transgenic plants -  
 XX  
 PS Claim 12; Page 45-49; 93pp; English.  
 XX  
 CC The present sequence is that of a portion of corn cellulose  
 CC synthase (CS), as deduced from an isolated cDNA clone (see AAY58265).  
 CC The invention relates to isolated CS nucleic acid fragments and to  
 CC CS polypeptides. It also relates to the construction of a chimeric  
 CC gene encoding all or a portion of the CS, in sense or antisense  
 CC orientation, where expression of the gene results in altered levels  
 CC of the CS in transformed host cells. The host cells can be used to  
 CC screen compounds for their ability to inhibit CS activity. CS  
 CC nucleic acids are also useful for producing transgenic plants having  
 CC altered levels of CS, and hence altered levels of fibre. CS may  
 CC also serve as a target for the development of novel herbicides.  
 XX  
 SQ Sequence 1165 AA;  
 Query Match 88.3%; Score 5228; DB 21; Length 1165;  
 Best Local Similarity 88.7%; Pred. No. 0;  
 Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;  
 Qy 1 MEASAGLVAGSHNNELVYTRD--GDPKRPPEBNGQVQCICGDDVGLAPGDPVVA 57  
 Db 72 MEASAGLVAGSHNNELVYTRDREBSGAAGAAARAP-CQICGVEVGFGDPEVVA 130  
 Qy 58 CNECAFVCRDCEYEREGTONCPCKTRFKRLKGCORVYGDDEBDVDLDFEFPN-MD 116  
 Db 131 CNECAFVCRDCEYEREGSQAQPCQCKTRFKRLKGCORVYGDDEBDVDLDFEFPN 190  
 Qy 117 G----HDSQSAESMLYHMSYGRGDPNGAPQAFOLNPNVPLITNGQVNDIPPEQHAL 172  
 Db 191 GAHEDDDQYVAESLRLAQNMYGRGD---AHGFSVPVNPVPLITNGQVNDIPPEQHAL 247  
 Qy 173 VPSFV---GGGKRIHPLPYADPSLPVOPRSMNDSKDLAAYGYSVAKEMEMKQROE 229  
 Db 248 VPSYVSGGGGKRIHPLPYADPNLPVOPRSMNDSKDLAAYGYSVAKEMEMKQROE 307  
 Qy 230 RMHQGNDGGD-DGDDADLPLMDEARQOLSRKILPSPSOINPYMTIIRLVYLGFFPH 288  
 Db 308 RLQHYRSEGGGDGDDADLPLMDEARQOLSRKILPSPSOINPYMTIIRLVYLGFFPH 367  
 Qy 289 YRWHPVNDAPALMLISVICIWFAMSWILDOFPKMLPIERTTYLDRLSLRDKGQSQ 348  
 Db 368 YRWHPADAFALMLISVICIWFAMSWILDOFPKMLPIERTTYLDRLSLRDKGQSQ 427  
 Qy 349 LAPIDFVSTVDPLKEPPLVTTNTVLSILSVDPYDKVSCVYSDGAAMLTFEALSETSE 408  
 Db 428 LAPIDFVSTVDPTKEPPLVTANTVLSILSVDPYDKVSCVYSDGAAMLTFEALSETSE 487  
 Qy 409 FAKKVPFCKRYNIEPRAPENYFOOKIDYLDKQVAAANFRRERAMKREYEERKYNALV 468  
 Db 488 FAKKVPFCKRYNIEPRAPENYFOOKIDYLDKQVAAANFRRERAMKREYEERKYNALV 547  
 Qy 469 AKAQVPEEGMTMOGTPMGPANNVBDHPGMIQVFLGQSGGRLCEGNEELRLVYVREKRP 528  
 Db 548 AKAQVPEEGMTMOGTPMGPANNVBDHPGMIQVFLGQSGGRLCEGNEELRLVYVREKRP 607  
 Qy 529 GYNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDLVGKVCY 588

|||||  
Db 608 GYNHKKAGANNALVRVSAVLNSNAAYLLNLDCDHYINNSKAIKRAMCFMMDPLVGKKVCY 667  
Qy 589 VQFQRDGDIDRHRVYANRVVFPDINKGIDGIGQPIYVTCGVFRQALYGDAPKTK 648  
Db 668 VQFQRDGDIDKNDRYANRVVFPDINKGIDGIGQPIYVTCGVFRQALYGDAPKTK 727  
Qy 649 KPSRTNCMPKWCFCGCCFGRNKKOKTKTKPTEKKKLLPFKKEENOSPALAGEIDEAA 708  
Db 728 KPSRTNCMPKWCFCGCCFGRNKKOKTKTKPTEKKKLLPFKKEENOSPALAGEIDEAA 787  
Qy 709 PGAENKAGIVNOCKLEKKFGQSSVFTSTLLENGTLKSPASPLKKAITHVISCYED 768  
Db 788 PGADIEKAGIVNOCKLEKKFGQSSVFTSTLLENGTLKSPASPLKKAITHVISCYED 847  
Qy 769 KTDWKEIGWYIGSVTEIDILTGFMHCHGMRSYICIPKRVAFKSPAPLNTSDRLHOVLRW 828  
Db 848 KTDWKEIGWYIGSVTEIDILTGFMHCHGMRSYICIPKRVAFKSPAPLNTSDRLHOVLRW 907  
Qy 829 ALGSIETFPNSHCELMYGGGGLKFLERFSYINSIYVPMWTSIPLLAYCTLPALICLTGKF 888  
Db 908 ALGSIETFPNSHCELMYGGGGLKFLERFSYINSIYVPMWTSIPLLAYCTLPALICLTGKF 967  
Qy 889 ITPELANNVASIMFMSLFCIFATSILEMWMSGVIGIDWWRNEQFVWIGVSSHFAVFOG 948  
Db 968 ITPELTNVASIMFMAIFCISVTGILEMRMSGVAIDWWRNEQFVWIGVSAHLFAVFOG 1027  
Qy 949 LKKTIAVDTSFTYTSKGDDEESELTYTFKWTLLIPPTLLLNPIGVAVGSAINNN 1008  
Db 1028 LKKTIAVDTSFTYTSKGDDEESELTYTFKWTLLIPPTLLLNPIGVAVGSAINNN 1087  
Qy 1009 GYESMGPLFGKLFPAFWYIYHLYPFLKGLVGRONRPTIYVWMSILLASIFSLMWADP 1068  
Db 1088 GYESMGPLFGKLFPAFWYIYHLYPFLKGLVGRONRPTIYVWMSILLASIFSLMWADP 1147  
Qy 1069 FLAKDGPPLLECCGLDCN 1086  
Db 1148 FLAKDGPPLLECCGLDCN 1165  
RESULT 9  
ID ABB93304 standard; Protein; 1084 AA.  
XX ABB93304;  
AC ABB93304;  
XX 31-MAY-2002 (first entry)  
DT 31-MAY-2002 (first entry)  
XX 31-MAY-2002 (first entry)  
DE Herbicidally active polypeptide SEQ ID NO 2515.  
XX Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
XX WO200210210-A2.  
XX 07-FEB-2002.  
XX 28-AUG-2001; 2001WO-EP09892.  
XX 28-AUG-2001; 2001WO-EP09892.  
XX (FARB) BAYER AG.  
XX Tietjen K. Weidner M;  
XX WPI; 2002-269010/31.  
XX Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms -

PS Claim 5; SEQ ID NO 2515; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins  
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.  
XX  
SQ Sequence 1084 AA;  
Query Match 72.3%; Score 4278.5; DB 23; Length 1084;  
Best Local Similarity 71.3%; Pred. No. 0;  
Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;  
Qy 1 MEASAGLVAGSHRNELVIRRDGDDPQKPRBRONGVQICGDDVGLAPGDPFVACNE 60  
Db 1 MNTGGRILAGSHRNELVIRRDGDDPQKPRBRONGVQICGDDVGLAPGDPFVACNE 60  
Qy 61 CAPPVCRDCYERREGEONCPQCKTRYKRLKCCORVTG--DEBEDGVDDLNERMMDGHD 119  
Db 61 CAPPVCRDCYERREGEONCPQCKTRYKRLKCCORVTG--DEBEDGVDDLNERMMDGHD 119  
Qy 120 SOSVAESMLYGHMSYGRGSDPNAGAPQAFQALNPVPLITNGQVDDIPPEGHALVPSFMG 179  
Db 120 PEHAABEALSSRLNTGSGIDSAAP-----GSGIPLITTCDEDDADWSDSHALIVPSTG 174  
Qy 180 GGRIRHPLFPADPSLPQPSMPSKDLAAYGVGVAMKERMENMKORO-ERM-----HOT 234  
Db 175 YGNRVYAPFTDSSAPQASMTVPQKDIAIYIGSVAMKORMEYWKROQEKQVITHEG 234  
Qy 235 GNDGGG---DDGDDADLPIMDEARQOLSRKILPSSQINPYEMIIIRLVVGFPEHYR 290  
Db 235 GNNRGSGNDDDELDDPMPWMDGROPLSRKILPSSQINPYEMIIIRLVVGFPEHYR 294  
Qy 291 VMHVPNDAPALMLISVCEIWFAMSWTLDPPKMPPIERTTYDRLSLRFDKSGQPSOLA 350  
Db 295 ILHPVNDAYGLWLTSVCEIWFAMSWTLDPPKMPPIERTTYDRLSLRFDKSGQPSOLA 354  
Qy 351 PIPFVSTVPLKBPPLVTNTNTVLSISVYPPDYKSCVYSDGAAMLTTEALSETSEFA 410  
Db 355 PVDVFTVDPLEBPPLVTNTNTVLSISVYPPDYKSCVYSDGAAMLTTEALSETSEFA 414  
Qy 411 KKWPFCRYNIEBRAPEWTFQOKIDYLDKQVIANFVREBRAMKREYEERFVRINALVAK 470  
Db 415 KKWVPECKKFNIEBRAPEWTFQOKIDYLDKQVIANFVREBRAMKREYEERFVRINALVAT 474  
Qy 471 AOKYVEEGMTMODGTFPGNNVRDHPGMIQVFLGHSQVDRDTEGNEIPRLVYVSREKPG 530  
Db 475 AOKYVEEGMTMODGTFPGNNVRDHPGMIQVFLGHSQVDRDTEGNEIPRLVYVSREKPG 534  
Qy 531 NHHKKAGANNALVRVSAVLTNAPYLNLDCDHYINNSKAIKRAMCFMMDPLVGKKVCY 590  
Db 535 DHHKKAGANNALVRVSAVLTNAPYLNLDCDHYINNSKAIKRAMCFMMDPLVGKKVCY 594  
Qy 591 FQRFQDIDRHRVYANRVVFPDINKGIDGIGQPIYVTCGVFRQALYGDAPKTKP 650  
Db 595 FQRFQDIDRHRVYANRVVFPDINKGIDGIGQPIYVTCGVFRQALYGDAPKTKP 654  
Qy 651 PSRTNCMPKWCFCGCCFGRNKKOKTKTKPTEKKKLLPFKKEENOSPALAGEIDEA--A 708  
Db 655 PSRTNCMPKWCFCGCCFGRNKKOKTKTKPTEKKKLLPFKKEENOSPALAGEIDEA--A 706  
Qy 709 PGAENKAGIVNOCKLEKKFGQSSVFTSTLLENGTLKSPASPLKKAITHVISCYED 768  
Db 707 PVSNEKRSATQLEKKFGQSSVFTSTLLENGTLKSPASPLKKAITHVISCYED 766  
Qy 769 KTDWKEIGWYIGSVTEIDILTGFMHCHGMRSYICIPKRVAFKSPAPLNTSDRLHOVLRW 828  
Db 767 KTDWKEIGWYIGSVTEIDILTGFMHCHGMRSYICIPKRVAFKSPAPLNTSDRLHOVLRW 826

QY 829 ALGSEIFPSNHCPLWYGGGLKFLERPSYINSIVYPTWTSIPLLAYCTLPALCLLTGKF 888  
 DB 827 ALGSEIFPSNHCPLWYGGGLKFLERPSYINSIVYPTWTSIPLLAYCTLPALCLLTGKF 886  
 QY 889 ITPELNVAISLPMFSLFICFATSLERMSVGVGIDDMWRNBOQFVIGVSHLPFAVROG 948  
 DB 887 ITPELNVAISLPMFSLFICFATSLERMSVGVGIDDMWRNBOQFVIGVSHLPFAVROG 946  
 QY 949 LLKVLAVGVTSTFTVTSKSGDDEEPESELYTFKMTLLIPPTLLLPFLLFVAVGVSNAINN 1008  
 DB 947 LLKVLAVGVTSTFTVTSKSGDDEEPESELYTFKMTLLIPPTLLLPFLLFVAVGVSNAINN 1006  
 QY 1009 GYSEKGRPLFGKLFPAFWIYVHLVPLKGLVGRONRPTTVIVWSLILASISLSLWRIIDP 1068  
 DB 1007 GYDSWGPLFGRLLFPAFWIYVHLVPLKGLVGRONRPTTVIVWSLILASISLSLWRIIDP 1066  
 QY 1069 FLAKDGPFLBEGGLDC 1085  
 DB 1067 FLAKDGPFLBEGGLDC 1082

## RESULT 10

ABB93949

ID ABB93949 standard: Protein; 1084 AA.

AC ABB93949;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3160.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB ) BAYER AG.

PI Tiejien K, Weidler M;

DR WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,

PT comparing aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

PS Claim 5; SEQ ID NO 3160; 261pp + Sequence listing; English.

CC The invention relates to identifying target proteins

CC (ABB90790-ABB94016) for herbicidally active compounds, comprising

CC aligning and comparing nucleic acid or amino acid sequences from plant

CC with nucleic acid or amino acid sequences from non-plant organisms using

CC suitable search parameters, where plant sequences having an E-value

CC greater by a factor of 3 than the E-value of most similar non-plant

CC sequences are selected. The polypeptides or nucleic acids encoding them

CC are useful for identifying modulators. The identified modulators are

CC useful as herbicides.

SQ Sequence 1084 AA;

Query Match 72.0%; Score 4262; DB 23; Length 1084;  
 Best Local Similarity 71.1%; Pred. No. 0;  
 Matches 778; Conservative 126; Mismatches 170; Indels 20; Gaps 10;

QY 1 MEASAGLVASHNRNELVYRRDDEGPKPREQNGVCGICGDDVGLACGDPFVACNE 60  
 DB 1 MNTGGRILIASHSHNNEFVLINADENARIRSVQELSGQTCICRDEIELTYDGEFPVACNE 60  
 QY 61 CAFFVPCRCYERERECTONCPOCKTRYKYLKGCORATGDEEBEGVDDLNEFWM--DGH 118  
 DB 61 CAFFVPCRCYERERECTONCPOCKTRYKYLKGCORATGDEEBEGVDDLNEFWM--DGH 120  
 QY 119 DSQSVASMLYGHMSYG--RGDDPGABOAFQOLNPVVELTNGQVDDIPEQHAL--VPSF 176  
 DB 121 GFDQVSEGMISIRNSGFPQSDLDASAPGSG---IFLLTYGDEVDVSISSDRALLVPS 176  
 QY 177 MGGGGRIRHPPLVADPSLPVQPSMDPSKDLAAYGVSVAKERMENWKOR--ERHQTC 235  
 DB 177 LGGNGNVHPVPSLSDPTVAAPRPMVYQKDLAAYGVSVAKERMENWKOR--ERHQTC 236  
 QY 236 NDGGD--DGGDADPLMDEARQOLSKILPLPSSQIPYPMIIRLVYGFPHYVWM 293  
 DB 237 HEGDPRDEDDGDADFPMDRGQPLSKIRIKSKINPYMILVRLVILGLFFHYRILH 296  
 QY 294 PVNDALMLISVICIEMFAMSWILDQPKMFPPIERETYLRLSLRDEKSGPSQLAPID 353  
 DB 297 PVKDAVALMLISVICIEMFAMSWILDQPKMFPPIERETYLRLSLRDEKSGPSQLAPID 356  
 QY 354 FVSTVDPPLKEPPLVTNTVLSLSVDPYDKVSCYSDGAMLTFEALSETSEPKKM 413  
 DB 357 VFSVTDPLKEPPLITNTVLSLAVDYPDVKVACVSDGAMLTFEALSETSEPKKM 416  
 QY 414 VPFCKRYNIEPRAPEMFQOKIDYLDKQVNAFPRERAKMREBEKRVINLVAAQK 473  
 DB 417 VPFCKRYNIEPRAPEMFQOKIDYLDKQVNAFPRERAKMREBEKRVINLVAAQK 476  
 QY 474 VPBEGMTMODGTPWPNVNRDHPGMIQVFLGSGGLDCEGNEPLRLVYVSEKRPYNNH 533  
 DB 477 VPBEGMTMODGTPWPNVNRDHPGMIQVFLGSGGLDCEGNEPLRLVYVSEKRPYNNH 536  
 QY 534 KKAGANALYRVSAVLTNAPYLLNLDCHYIINNSKAIKEAMCFMMDPLGKRCVYPOQ 593  
 DB 537 KKAGANALYRVSAVLTNAPYLLNLDCHYIINNSKAIKEAMCFMMDPLGKRCVYPOQ 596  
 QY 594 RFDGIDRHRANNNVFFDINMKGLDGIQPIYVGVCFRRQALGVYAPTKKPPSR 653  
 DB 597 RFDGIDRHRANNNVFFDINMKGLDGIQPIYVGVCFRRQALGVYAPTKKPPSR 656  
 QY 654 TCNCPMKCFCCCFGRKOKKTKPKTEKKKLFPKKEENOSPAVALGEIDE--AAPGA 711  
 DB 657 TCNCPMKCFCCCFGRKOKKTKPKTEKKKLFPKKEENOSPAVALGEIDE--AAPGA 710  
 QY 712 ENERAGIVNOOKLEKKKGQSSVFTSTLNGGTLLKASPAISLKEAIIHIVISGEYEDKT 771  
 DB 711 NVEQSTAMQKLEKKKGQSSVFTSTLNGGTLLKASPAISLKEAIIHIVISGEYEDKT 770  
 QY 772 MGKEIGMIVGSVEDILTGFMCHGMRISYICIPKRAVFGSAPLNSDLHOLVRLMALG 831  
 DB 771 MGKEIGMIVGSVEDILTGFMCHGMRISYICIPKRAVFGSAPLNSDLHOLVRLMALG 830  
 QY 832 SIEIFPSNHCPLWYGGGLKFLERPSYINSIVYPTWTSIPLLAYCTLPALCLLTGKFTP 891  
 DB 831 SIEIFPSNHCPLWYGGGLKFLERPSYINSIVYPTWTSIPLLAYCTLPALCLLTGKFTP 890  
 QY 892 ELNNAVASLPMFSLFICFATSLERMSVGVGIDDMWRNBOQFVIGVSHLPFAVROGLK 951  
 DB 891 ELNNAVASLPMFSLFICFATSLERMSVGVGIDDMWRNBOQFVIGVSHLPFAVROGLK 950  
 QY 952 VIAGVDNFTVTSKSGDDEEPESELYTFKMTLLIPPTLLLPFLLFVAVGVSNAINN 1011  
 DB 951 VIAGVDNFTVTSKSGDDEEPESELYTFKMTLLIPPTLLLPFLLFVAVGVSNAINN 1010  
 QY 1012 SWGPLFGKLFPAFWIYVHLVPLKGLVGRONRPTTVIVWSLILASISLSLWRIIDPFLA 1071  
 DB 1011 SWGPLFGKLFPAFWIYVHLVPLKGLVGRONRPTTVIVWSLILASISLSLWRIIDPFLA 1070  
 QY 1072 KDDGPFLBEGGLDC 1085

Db | ||| ||| ||| |||  
1071 K-GGFILEICGLDC 1083

RESULT 11  
AAW33818  
ID AAW33818 standard; Protein; 1084 AA.

XX AAW33818;

DT 06-JUL-1998 (first entry)

DE Arabidopsis cellulose biosynthetic protein Ath-A.

KM Cellulose; cellulose synthase; RSM1 protein; beta-1,4-glucan;

XX transgenic plant.

OS Arabidopsis thaliana var. Columbia.

PN W09800549-A1.

XX 08-JAN-1998.

PF 24-JUN-1997; 97WO-AU00402.

PR 27-JUN-1996; 96AU-0000699.

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.

PI (AUSU ) UNIV AUSTRALIAN NAT.

PI Artold A, Betzner AS, Peng L, Williamson RE;

DR WPI: 1998-086974/08.

DR N-PSDB; AAW06566.

PT DNA encoding cellulose biosynthetic enzyme - useful for manipulation

XX of cellulose and beta-1,4-glucan

PS Claim 29; Page 130-135; 207pp; English.

XX This polypeptide is encoded by cDNA clone Ath-A (see AAW06566)

CC that was isolated from an Arabidopsis thaliana cDNA library using

CC PCR primers (see AAT9632-34) based on cellulose synthase RSM1

CC genomic clone 23H12 (see AAW06563) and EST clone AAT20782 (see

CC AAW06562). It is closely related to Arabidopsis cellulose synthase RSM1

CC protein (see AAW33816-17). Claimed nucleic acid molecules (see

CC AAW06562-69) coding for claimed polypeptides (see AAW33816-20 and

CC AAW46202) involved in cellulose biosynthesis can be used to

CC manipulate the cellulose and/or beta-glucan content of transgenic

CC plants. Expression of nucleic acids in the sense orientation

CC increases the level of cellulose and reduces the level of

CC non-crystalline beta-1,4-glucan and starch, providing plants with

CC modified strength and/or shape and/or fibre properties, or having

CC increased resistance to stresses or pests. Antisense, ribozyme or

CC co-suppression molecules can be used to reduce the cellulose

CC content of a transgenic plant, e.g. to improve digestibility or to

CC alter carbon partitioning such that increased carbon is available

CC for growth, rather than deposited as cellulose.

XX Sequence 1084 AA;

Query Match 71.8%; Score 4250.5; DB 19; Length 1084;

Best Local Similarity 70.8%; Pred. No. 0;

Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;

QY 1 MEASAGLVAGSHNNELVIRDDGPGKPRRQNGGVCOICGPDVGLAPGDFVACNE 60

Db 1 MNTGGRLLAGSHNNEFLVIRDDGPGKPRRQNGGVCOICGPDVGLAPGDFVACNE 60

QY 61 CAFVPCDCYERREBEGTNCPOCKTRYKRLKGCQRYTG--DEEDGDVLDLNEFNMDGH 119

Db 61 CAFVPCDCYERREBEGTNCPOCKTRYKRLKGCQRYTG--DEEDGDVLDLNEFNMDGH 119

QY 120 SQSVASMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWVDIPRBOHALVPSFMG 179

Db 120 PERHAERALSRLKATRGGLDSAP-----GSGIPLITTYCDEBDADKSDHALVPPSTG 174

QY 180 GGRRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKEMENKQRC-ERM-----HQT 234

Db 175 YGNRVYAPFPTDSAPFQASWVPQDIARYGYGSVAMKQRMVWRQGEKQVITHEG 234

QY 235 GNDGGG---DDGDDADLPMDARQOLSKITLPSSQINPYMIIIRLVIGFFPHYR 230

Db 235 GNNRGSGNDDDELDDPMDMDGRQPLSKPLRSRIRIPYMLILCRALAIIGLFFHYR 234

QY 291 VMHPVNDAPALMTISVCEIWPAMSWITLDPFPMPIREERTYLDRLSRBDKXGQSQLA 350

Db 295 ILHPVNDAYGLMLTSVCEIWPAMSWITLDPFPMPIREERTYLDRLSRBDKXGQSQLA 354

QY 351 PIDFVSTVDPLKEPPLVTNTVLSILSVYPVVKYSCVYSDGAAMLTFEALSETSEFA 410

Db 355 PVDVFSVSTVDPLKEPPLVTNTVLSILSVYPVVKYSCVYSDGAAMLTFEALSETSEFA 414

QY 411 KKRVPCCKRYNIEPRAPWTFOQKIDYLDKQVAAVFRERANKREYEERKVINMLVAK 470

Db 415 TKRVPCCKRYNIEPRAPWTFOQKIDYLDKQVAAVFRERANKREYEERKVINMLVAK 474

QY 471 AOKVPERGWTMODGTPMGNVNRDHPGMOVFLGSGGGLDCBENELPRLVYGRERKPGY 530

Db 475 AOKVPERGWTMODGTPMGNVNRDHPGMOVFLGSGGGLDCBENELPRLVYGRERKPGY 534

QY 531 NHHKAGAMNALYRVSAYLTNAAYLLNDCDHYINNSKAIKAMCFMMDPLGKVCYQ 590

Db 535 DHHKAGAMNSLIRVSAYLTNAAYLLNDCDHYINNSKAIKAMCFMMDPLGKVCYQ 594

QY 591 FPOPFQIDHNDRYANRVVFPDINKGLDGIQPIYVGTGCVPRQALYGYDAPTKPK 650

Db 595 FPOPFQIDHNDRYANRVVFPDINKGLDGIQPIYVGTGCVPRQALYGYDAPTKPK 654

QY 651 PSRTCNCFKMCFCGCCFGNRKOKT--TKPKTEKKKLEFFKEENQSPAYALGEIDA- 707

Db 655 PSRTCNCFKMCFCGCCFGNRKOKT--TKPKTEKKKLEFFKEENQSPAYALGEIDA- 704

QY 708 -APGAEKAKGIYNOQLEKKFGQSSVFTVSTLLENGGTLKSPASILKEALHVISG 766

Db 705 IYVSNVNEKSEANTOLKEKKFGQSPVFAVANLQNGGVNRNAPCLRLBALQVISG 764

QY 767 EBDTNGKEIGMITYSVTEIDILTFPMHCHGMSYICIPRVAFKGSAPLNLSDRLH 826

Db 765 QDRTMGKEIGMITYSVTEIDILTFPMHCHGMSYICIPRVAFKGSAPLNLSDRLH 824

QY 827 RMLGSIETIFFSNHCPLMYGGGLKFLERFSYINSLVYPTWISPLIAYCTTLPAICLT 886

Db 825 RMLGSIETIFFSNHCPLMYGGGLKFLERFSYINSLVYPTWISPLIAYCTTLPAICLT 884

QY 887 KFTPELNVANYSIMFSLFICIPATSLILEMRSVGVDIMWRREOFVIGVSHLPAV 946

Db 885 KFTPELNVANYSIMFSLFICIPATSLILEMRSVGVDIMWRREOFVIGVSHLPAV 944

QY 947 QGLKVIAGVDTSTFTYSKGGDEEFSSELYTFKMTLLIPPTLLLNFGVAVGSNAI 1006

Db 945 QGLKVIAGVDTSTFTYSKGGDEEFSSELYTFKMTLLIPPTLLLNFGVAVGSNAI 1004

QY 1007 NNGVSGPLFGLTFPAFWIVLTPPLKLVORQRNPTPIVYVMSGLASISLWRI 1066

Db 1005 NNGVSGPLFGLTFPAFWIVLTPPLKLVORQRNPTPIVYVMSGLASISLWRI 1064

QY 1067 DPEFLAKDDGPLEEGGLDC 1085

Db 1065 DPEFLAKDDGPLEEGGLDC 1082

RESULT 12

ABB91831

XX ID ABB91831 standard; Protein; 1088 AA.

AC ABB91831;  
 XX  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Herbicidally active polypeptide SEQ ID NO 1042.  
 XX  
 KW Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN WO200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001WO-EP09892.  
 XX  
 PR 28-AUG-2001; 2001WO-EP09892.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Tietjen K, Weidler M;  
 XX  
 DR WPI, 2002-269010/31.  
 XX  
 PT Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms -  
 XX  
 PS Claim 5; SEQ ID NO 1042; 261pp + Sequence Listing; English.  
 XX  
 CC The invention relates to identifying target proteins  
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
 CC aligning and comparing nucleic acid or amino acid sequences from plant  
 CC with nucleic acid or amino acid sequences from non-plant organisms using  
 CC suitable search parameters, where plant sequences having an E-value  
 CC greater by a factor of 3 than the E-value of most similar non-plant  
 CC sequences are selected. The polypeptides or nucleic acids encoding them  
 CC are useful for identifying modulators. The identified modulators are  
 CC useful as herbicides.  
 CC  
 XX  
 SQ Sequence 1068 AA;  
 Query Match 71.1%; Score 4209.5; DB 23; Length 1088;  
 Best Local Similarity 70.2%; Pred. No. 0;  
 Matches 769; Conservative 127; Mismatches 179; Indels 21; Gaps 10;

DB 359 APVDVFSVTDPLKEPPLITANTVLSTLAVDYVEKVAQYSDGGAAMLTFFALSTAEF 418  
 QY 410 AKKVVPECKRKNIEPRAPWYFOOKIDYLDKDKVAANFRERRAMKEEYEFKRYINALVA 469  
 DB 419 ARKKVPCKKFSIEPRAPWYFSGKMDYLGHKDPDAFVWRBAMKDYEEFKIKYNLVS 478  
 QY 470 KAKVPEEGWTMODGTWPGANNVVDHPGMIQVELGSGGGLDCEGNELPRLVYVREKRP 529  
 DB 479 VSKVPEDEGTWMDGTWPGANNVVDHPGMIQVELGSHGVCMDGNELPRLVYVREKRP 538  
 QY 530 YNHKKAGANALVRSAYLTNAPELLNDCHYINNSKAIKRAMCMMDPLLGKRCYV 589  
 DB 539 FDHKKAGANSLIRSAVLSNAPYLNVDCDHYINNSKAIKRAMCMMDPOSGKICVY 598  
 QY 590 QPQRFQIDRHRDPRANRVVFPDINMKGLDGIQPIYVGTGCVFRQALYGYDAPKTK 649  
 DB 599 QPQRFQIDRHRDPRANRVVFPDINMKGLDGIQPIYVGTGCVFRQALYGYDAPKTK 658  
 QY 650 PPSRTCCMPKWCPCCCCFGNRRKQKTTKPKTEKKLLFPKKEBNSPAYALGEIDBAAP 709  
 DB 659 PGRTCNCPKWC-CGCC-GMRK-KTGKVKDQNRK---KPKETSKQIHAIHEIGLQ 711  
 QY 710 GAENEKAGYNQOKLEKKFGOSSVFTSTLTLENGCTIKASPSILKEATHVTSCEYEDK 769  
 DB 712 VTNAENNSETAQLERKFFQSPVLVASTLLNGVPSNVNPAFLRESIQVISGYESK 771  
 QY 770 TDMCKEIGMTYGSVTEIDLTGPKMKCHGWSRISYICIPKRVAFKGSAPLNTDRLHQVLRMA 829  
 DB 772 TEMKEIGMTYGSVTEIDLTGPKMKCHGWSRISYICIPKRVAFKGSAPLNTDRLHQVLRMA 831  
 QY 830 LGSIEIFPSNHCPILWYGGGLKFLERFSYINSIIVPWTISIPLLAICTLPALCTLNGKFI 889  
 DB 832 LGSVEILSRHCPILWYGGGLKFLERFSYINSIIVPWTISLPLVYCSLPALCTLNGKFI 891  
 QY 890 TPELNANVASLWMSLFCIFATSTLEMRSGVIGIDMWREQWVIGVSSHLPAYVQGL 949  
 DB 892 VPEISNVAAGLILFLMFMSIIVTGLIEMQWKIGIDMWREQWVIGVSSHLPAYVQGL 951  
 QY 950 LKVIAGVDTFTVTSKGGDEEFSELYTFKMTTLLIPTLLLNLFIGVAVGASNAIING 1009  
 DB 952 LKVIAGVDTFTVTSKGGDEEFSELYTFKMTTLLIPTLLLNLFIGVAVGASNAIING 1011  
 QY 1010 YESWGPLFGKLFPAFWIVLHLYPFLKGLVGRQNRTPFIVVMSILLASIFSLMWRIDPF 1069  
 DB 1012 YDSWGPLFGKLFPAFWIVLHLYPFLKGLKQBRVPTIILVMSILLASIFSLMWRIDPF 1071  
 QY 1070 LAKDGPFLBECGLDC 1085  
 DB 1072 VSK-DGVLLETGGLDC 1086

RESULT 13  
 ABB93411  
 ID ABB93411 standard; Protein; 1069 AA.  
 XX  
 AC ABB93411;  
 XX  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Herbicidally active polypeptide SEQ ID NO 2622.  
 XX  
 KW Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN WO200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001WO-EP09892.  
 XX  
 PR 28-AUG-2001; 2001WO-EP09892.





CC PtcA and AtaxCela respectively. The polynucleotides may be introduced  
 CC into plants where they are useful for altering the growth of a plant.  
 CC Vectors comprising a cellulose synthase promoter, or a fragment of the  
 CC promoter containing one or more mechanical stress response elements  
 CC (MSRE), operably linked to a cellulose synthase gene are useful for  
 CC causing stress-induced gene expression in a plant cell. A cDNA encoding a  
 CC protein that binds to a positive MSRE may be delivered into the plant to  
 CC increase expression of cellulose synthase. A cDNA in an antisense  
 CC orientation may be delivered to reduce expression of cellulose synthase.  
 CC The polynucleotides enable the development of transgenic tree varieties  
 CC having increased cellulose content, decreased lignin content and  
 CC therefore improved wood fibre characteristics. The cellulose synthase  
 CC promoters are useful for synthesizing greater quantities of  
 CC high-crystalline cellulose in plant and preferably in trees, which  
 CC permits an increase in proportion of cellulose in transgenic plants,  
 CC greater strength of juvenile wood and fiber and acceleration of overall  
 CC growth rate.  
 CC Note: This nucleotide sequence is given in the specification as  
 CC SEQ ID NO: 5. SEQ ID NO: 5 is referred to in claim 8 but it is described  
 CC as a polypeptide sequence.

XX Sequence 1080 AA:

Query Match 70.7%; Score 4185.5; DB 22; Length 1080;  
 Best Local Similarity 70.7%; Pred. No. 0;  
 Matches 769; Conservative 126; Mismatches 172; Indels 21; Gaps 11;

QY 7 LVAGSHNRELVIYRRDGPGRPKPRREGQVQIQIGDDVGLAPGSDPFAVACNECAFPC 66  
 DB 4 LVAGSHNRELVIYRRDGPGRPKPRREGQVQIQIGDDVGLAPGSDPFAVACNECAFPC 63  
 QY 67 RDCYEREREGTQNPQCKTRRYKRLKGCORVGTGDEEEDVDLDNEFMN--DGHDSQSYA 124  
 DB 64 RDCYEREREGTQNPQCKTRRYKRLKGCORVGTGDEEEDVDLDNEFMN--DGHDSQSYA 123  
 QY 125 ESMLYGHMSYG-RGSDPNAPQAFQNLNPNVPLITNGQVDDIPEEHAL-VPSFMGGGSK 182  
 DB 124 ESMYSIRSRNSGPPQSLDAPPSQ---IPLTYDEDEVEISSDHALIVPSPISGGHGN 179  
 QY 183 RIHPPLYPADPSLVPQSRMSPSKDLAAYGGSVAMERMMKKORO-BEMHQDTGGDGD 241  
 DB 180 RIHPPLYPADPSLVPQSRMSPSKDLAAYGGSVAMERMMKKORO-BEMHQDTGGDGD 239  
 QY 242 --DGDADADPLMDEARQQLSRKIPSPSQINPYRMIIIRLVYLGFFFHYVMHPVND 299  
 DB 240 FEDGDDADPFPMDGEGQPLSMKIPKSSKINPYRMILVIRLVILGFFHYRILHPKXAY 299  
 QY 300 ALMLISVCEIFPANSWIIDQPKWPIRETYLDRSLRFDKEGQPSQLAPIDFFVSTV 359  
 DB 300 ALMLISVCEIFPANSWIIDQPKWPIRETYLDRSLRFDKEGQPSQLAPIDFFVSTV 359  
 QY 360 DPLKEPPLVITNTVLSILSDVYVUDKVCYVSDGGAAMLTPEALSTSEFAKKWVPFCR 419  
 DB 360 DPLKEPPLVITNTVLSILSDVYVUDKVCYVSDGGAAMLTPEALSTSEFAKKWVPFCR 419  
 QY 420 YNIEPRAPPEWFOQKIDYLDKVAANFVERRAMKREYEEFKVIRINATAKAKQKPEEGM 479  
 DB 420 YNIEPRAPPEWFOQKIDYLDKVAANFVERRAMKREYEEFKVIRINATAKAKQKPEEGM 479  
 QY 480 TQODGTTPWPGNNVRDHPGMIQVFLGSGGLDCEGNLRLVYVSREKRGYNNHKKAGM 539  
 DB 480 TQODGTTPWPGNNVRDHPGMIQVFLGSGGLDCEGNLRLVYVSREKRGYNNHKKAGM 539  
 QY 540 NALVRSAVLTNAPYLLNDCHYINNSKAKEAMCFMMDPLLGKVCYVQFPQRPDGD 599  
 DB 540 NALVRSAVLTNAPYLLNDCHYINNSKAKEAMCFMMDPLLGKVCYVQFPQRPDGD 599  
 QY 540 NSLIRSGVLSNAPYLLNDCHYINNSKAKEAMCFMMDPQSGKICIVQFPQRPDGD 599  
 DB 540 NSLIRSGVLSNAPYLLNDCHYINNSKAKEAMCFMMDPQSGKICIVQFPQRPDGD 599  
 QY 600 RHDYRANRVVFFDINMGKLDGIQGPVYVGTGCFVRROALYGVDAKTKKPPSRITCNP 659  
 DB 600 RHDYRANRVVFFDINMGKLDGIQGPVYVGTGCFVRROALYGVDAKTKKPPSRITCNP 659  
 QY 660 RHDYRANRVVFFDINMGKLDGIQGPVYVGTGCFVRROALYGVDAKTKKPPSRITCNP 658  
 DB 660 RHDYRANRVVFFDINMGKLDGIQGPVYVGTGCFVRROALYGVDAKTKKPPSRITCNP 658  
 QY 660 KNCFCGCCFGNRKQKTKTPKTEKKLLFFKKEENSPLVALGEIDEA--FGAENKAG 717  
 DB 660 KNCFCGCCFGNRKQKTKTPKTEKKLLFFKKEENSPLVALGEIDEA--FGAENKAG 717

DB 659 KWCIL-LCFSGRKNRKAKTVADKK---KNREASQIHALENIEEGRGHKVLNVEOST 712  
 QY 718 IVNQKLEKKRGQSSVFTSTLLNGLTKASAPSLKEAIIHISGVEDKTDMGKEIG 777  
 DB 713 EAMQKQKXQGSPPVPAASARLENCGMARASACILKEAIIQVINSIGYEKTEMGKEIG 772  
 QY 778 WYGSVTEIDILITGFMQCHGRSIIYCIKPRVAFPGSAPLNSIDLHOVLAMALGSTEIIF 837  
 DB 773 WYGSVTEIDILITGFMQCHGRSIIYCIKPRVAFPGSAPLNSIDLHOVLAMALGSTEIIF 832  
 QY 838 SNHCPIMWYGGGKIFLERFSYINSIYVPMWSTILLAYCTLPALCLITGKITTEELNVA 897  
 DB 833 SRHCPIMWYGGGKIFLERFSYINSIYVPMWSTILLAYCTLPALCLITGKITTEELNVA 892  
 QY 898 SLWMSLFCIFATSIEMRWGVIDDMRNBOFWYIGVSSHLFAVFOGLKVIYAGVD 957  
 DB 893 SLWMSLFCIFATSIEMRWGVIDDMRNBOFWYIGVSSHLFAVFOGLKVIYAGVD 952  
 QY 958 TSFTVTSKGGDDEPSELYTFKMTLLIPTTLILLNFIQVAVSNAINNGYESWGP 1017  
 DB 953 TSFTVTSKGGDDEPSELYTFKMTLLIPTTLILLNFIQVAVSNAINNGYESWGP 1012  
 QY 1018 GKLFPAPVIVHLYPFLKGLVGRNRPRTIYVASSILLASTFSLMWRIDPFLAKDQPL 1077  
 DB 1013 GRLFPALVILHLYPFLKGLKODRMPPTIIVMSILLASTILLMWRVNPFAK-GQPI 1071  
 QY 1078 LEECGLDG 1085  
 DB 1072 LEECGLDG 1079

RESULT 15  
 ID AAY58837 standard; Protein: 1039 AA.  
 DB AAY58837;  
 AC AAY58837;  
 XX 08-MAY-2000 (first entry)  
 DT  
 XX  
 DE Soybean cellulose synthase (full-length).  
 XX  
 KM Soybean; cellulose synthase; transgenic plant.  
 XX  
 OS Glycine max.  
 XX  
 FT Key Location/Qualifiers  
 FT Misc-difference 201 /note="encoded by AGAA (frameshift)"  
 FT  
 PN WO200004166-A2.  
 XX  
 PD 27-JAN-2000.  
 XX  
 PF 13-JUL-1999; 99WO-US15871.  
 XX  
 PR 14-JUL-1998; 98US-0092844.  
 XX  
 PI (DURO) DU PONT DE NEMOURS & CO E. I.  
 XX  
 PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G,  
 PI Rafalski JA, Thorpe CJ.  
 XX  
 DR WPI, 2000-182431/16.  
 XX  
 DR N-PsDB; AA258268.  
 XX  
 PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
 PT probes for isolating cDNAs and genes encoding homologous proteins, for  
 PS producing transgenic plants -  
 XX  
 PS Claim 6; Page 56-60; 93pp; English.  
 CC The present sequence is that of full-length soybean cellulose  
 CC synthase (CS), as deduced from an isolated cDNA clone (see AAY58268).  
 CC

CC The invention relates to isolated CS nucleic acid fragments and to  
 CC CS polypeptides. It also relates to the construction of a chimeric  
 CC gene encoding all or a portion of the CS, in sense or antisense  
 CC orientation, where expression of the gene results in altered levels  
 CC of the CS in transformed host cells. The host cells can be used to  
 CC screen compounds for their ability to inhibit CS activity. CS  
 CC nucleic acids are also useful for producing transgenic plants having  
 CC altered levels of CS, and hence altered levels of fibre. CS may  
 CC also serve as a target for the development of novel herbicides.

SQ Sequence 1039 AA;

Query Match 67.4%; Score 3986; DB 21; Length 1039;

Best Local Similarity 67.6%; Pred. No. 0; Mismatches 159; Indels 60; Gaps 14;

Matches 738; Conservative 135; Mismatches 159; Indels 60; Gaps 14;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPKPREQNGOVQOLGDDVGLAPGDDPVACNE 60  
 DB 1 MEASAGLVAGSHNRNELVIRRDGDPKPREQNGOVQOLGDDVGLAPGDDPVACNE 58  
 QY 61 CAFPVCRDCYERREBGTQNCPCQKTRRYRLGQCRVTGDEEEDVDLNEFNMDGHS 120  
 DB 59 CGFPVCRDCYERREBGTQNCPCQKTRRYRLGQCRVTGDEEEDVDLNEFNMDGHS 118  
 QY 121 Q--SVAESMLYGHMVGGRGDPNGAPQAFQALNPVPLTNGQ--MVDIPPEQHALVPS 175  
 DB 119 KHGQVAEAMLHGRMSYGRGDEDDNSQF--PTPVLAGRSRPSGSEFPISNAVGDQ 173  
 QY 176 FMGGG-GKRIRHPLPADPSLPVQPRSMPSKDLAAYGVGSVAMKEMENKORERHQT 234  
 DB 174 MLSSSLHGRVHPYPSERB--SARWDEKXD-----GKDRMDMKLQ----- 215  
 QY 235 GNDGGDDGDDADLPLMEARQQLSRKIPLPSSQINPYRMIIIRLVGLFFHYVWMP 294  
 DB 216 GNLGPEPD-EDDAMLDARQPLSRKVP IASSKINPYRMIVARLVILAFELRYRLMP 274  
 QY 235 VNDALALMLISYICGIMFAMSWILDOFPKMPFERETIDRLSLAFDKEGQSPQLAPIDF 354  
 DB 275 VHDALGLMLTSLICEIWFASFILDOFPKMPFERETIDRLSLIRYEREGEPMMLAPVDV 334  
 QY 355 FVSTVDPLEKPEPLVNTNTVLSLSDYVPYDKVSCVSDGAMLTFFBALSETSEFAKKV 414  
 DB 335 FVSTVDPLEKPEPLVNTNTVLSLSDYVPYDKVSCVSDGAMLTFFBALSETSEFAKKV 394  
 QY 415 PCKRNYIEPRAPEMWFOOKIDYLDKQVAAFVREERAMKREYEEFKVINALVAKQKV 474  
 DB 395 PCKKFSIEPRAPEMWFOOKIDYLDKQVPTFVKERAMKREYEEFKVINALVAKQKV 454  
 QY 475 PEEGWTMODGTWPGNNVADHFGMTQVFLGSGGLDCBGNELPRLVYVSREKRGGINHK 534  
 DB 455 POGGWMODGTWPGNNVADHFGMTQVFLGSGGLDCBGNELPRLVYVSREKRGGINHK 514  
 QY 535 KAGANALVRVASVLTNA.PYLLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYQOPOR 594  
 DB 515 KAGANALVRVASVLTNA.PYLLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYQOPOR 574  
 QY 595 FDGIDRHDRYANRVVFDINKKGLDGIQPIYVGTGCVFRQALXYGDAPTKKPSPRT 654  
 DB 575 FDGIDTHDRYANRVVFDINKKGLDGIQPIYVGTGCVFRQALXYGDAPTKKPSPRT 634  
 QY 655 CNWPKWCFCCCFGNRQOKTTPKTEKKLLFFKKEENOSPAYALGEIDEAPGAENE 714  
 DB 635 S-----CDCCPCFGSRKK-----YKERND---ANGBA--ASLKGMDDD 667  
 QY 715 KAGIYMOQLKKKFGQSSVFTVSTLLENGTLKSASPASLKEAHHVISCYEDKTDWCK 774  
 DB 668 KEVLSQNFERRKFGQSSVFTVSTLMEGVPSSSPALLKEAHHVISCYEDKTEWGL 727  
 QY 775 BIGMIYGVTEIDILGFKHCHGMRISYICIPKRVAFKSGAPLNTSDRLHOVLRNALGSI 834  
 DB 728 ELGMIYGVTEIDILGFKHCHGMRISYICIPKRVAFKSGAPLNTSDRLHOVLRNALGSI 787  
 QY 835 IFFSNHCPLMVGY-GGGLKFLERFSYINSIVYPMTSIPLLAYCTLPAICLLTGKFTPEL 893

DB 788 IFFSHHCPLMVGYGFKKELKMLERFAVANTTYPTTSIPLVAYCILPAVCILTDKFTPEPI 847  
 QY 894 NNVASLWFMSLIFICIFATSIEMKMSGVIGIDWNRNQFWVIGVSSHLPAVFOGLKVI 953  
 DB 848 STPAGLVFVALFSSIIATGILELMKSGVISEMWRNQFWVIGVSSHLPAVFOGLKVI 907  
 QY 954 AGVTSPTVYNSKGGDDDEFSRLYFKWTTLLIPTTLLLNLFQVAVAGVSNAINNGYSW 1013  
 DB 908 AGIDTNPFTVNSKATDDEFGELYTFKWTLLIPTTLLIINIVGVVAGISDAINNGYSW 967  
 QY 1014 GPLFGKLFPAFVIVHLYPFLKGLVGRONRPTIVIVMSIILASIFSLMVRIDPELAKD 1073  
 DB 968 GPLFGKLFPSFVIVHLYPFLKGLMGRONRPTIVIVMSIILASIFSLMVRIDPELAKT 1027  
 QY 1074 DGPLLEEGGLDC 1085  
 DB 1028 KGPDTKLGINC 1039

Search completed: December 15, 2003, 13:02:49  
 Job time : 51 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:04:25 ; Search time 36 Seconds

(without alignments)  
5610.507 Million cell updates/sec

Title: US-09-720-383c-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNRRLVIR...DPLAKDDGPIELGCGLDGN 1086

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

684280

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5918	100.0	1086	10	US-09-900-237-10 Sequence 10, Appl
2	5918	100.0	1086	12	US-10-209-059-50 Sequence 10, Appl
3	5918	100.0	1086	12	US-10-160-719-18 Sequence 18, Appl
4	5918	100.0	1086	12	US-10-160-719-38 Sequence 38, Appl
5	5570.5	94.1	1148	10	US-09-900-237-4 Sequence 4, Appl
6	5408.5	91.4	1059	12	US-10-209-059-42 Sequence 42, Appl
7	5408.5	91.4	1059	12	US-10-160-719-2 Sequence 2, Appl
8	5228	88.3	1084	12	US-10-209-059-18 Sequence 18, Appl
9	5228	88.3	1094	12	US-10-160-719-26 Sequence 26, Appl
10	5228	88.3	1094	12	US-10-160-719-46 Sequence 46, Appl
11	5228	88.3	1165	10	US-09-900-237-8 Sequence 8, Appl
12	4250.5	71.8	1084	15	US-10-229-193-8 Sequence 8, Appl
13	3991	67.4	1075	12	US-10-209-059-2 Sequence 2, Appl
14	3991	67.4	1075	12	US-10-160-719-10 Sequence 10, Appl
15	3991	67.4	1075	12	US-10-160-719-34 Sequence 34, Appl



16	3991	67.4	1075	12	US-10-160-719-54 Sequence 54, Appl
17	3986	67.4	1039	10	US-09-900-237-14 Sequence 14, Appl
18	3977.5	67.2	1074	12	US-10-209-059-46 Sequence 46, Appl
19	3977.5	67.2	1074	12	US-10-160-719-14 Sequence 14, Appl
20	3977.5	67.2	1074	12	US-10-160-719-22 Sequence 22, Appl
21	3977.5	67.2	1074	12	US-10-160-719-42 Sequence 42, Appl
22	3970.5	67.1	1091	10	US-09-900-237-26 Sequence 26, Appl
23	3945.5	66.7	1081	15	US-10-229-193-6 Sequence 6, Appl
24	3945.5	66.6	1081	15	US-10-229-193-12 Sequence 12, Appl
25	3917	66.2	1076	12	US-10-209-059-14 Sequence 14, Appl
26	3917	66.2	1076	12	US-10-160-719-58 Sequence 58, Appl
27	3905.5	66.0	1077	12	US-10-209-059-10 Sequence 10, Appl
28	3905.5	66.0	1077	12	US-10-160-719-6 Sequence 6, Appl
29	3905.5	66.0	1077	12	US-10-160-719-30 Sequence 30, Appl
30	3801	65.9	1078	12	US-10-209-059-26 Sequence 26, Appl
31	3880.5	65.6	1079	12	US-10-267-459-6 Sequence 6, Appl
32	3880.5	65.6	1079	12	US-10-209-059-22 Sequence 22, Appl
33	3876.5	65.5	1065	15	US-10-229-193-10 Sequence 10, Appl
34	3875.5	65.5	1065	10	US-09-900-237-33 Sequence 33, Appl
35	3818	64.5	1080	10	US-09-900-237-30 Sequence 30, Appl
36	3786.5	64.0	1043	12	US-10-160-719-50 Sequence 50, Appl
37	3611.5	61.0	927	12	US-10-209-059-30 Sequence 30, Appl
38	3552	60.0	881	10	US-09-838-539-8 Sequence 8, Appl
39	3521	59.5	793	10	US-09-900-237-18 Sequence 18, Appl
40	3469.5	58.6	974	10	US-09-838-539-6 Sequence 6, Appl
41	3442	58.2	1007	12	US-10-209-059-28 Sequence 28, Appl
42	3318	56.1	821	12	US-10-267-459-2 Sequence 2, Appl
43	3318	56.1	821	12	US-10-209-059-6 Sequence 6, Appl
44	3261.5	55.1	740	10	US-09-900-237-24 Sequence 24, Appl
45	2956.5	50.0	701	10	US-09-900-237-32 Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-09-900-237-10  
Sequence 10, Application US/09900237  
Patent No. US20020120124A1  
GENERAL INFORMATION:  
APPLICANT: Allen, Stephen  
TITLE OF INVENTION: Plant Cellulose Synthases  
FILE REFERENCE: B01170 US CIP  
CURRENT APPLICATION NUMBER: US/09/900,237  
PRIOR FILING DATE: 2001-07-06  
PRIOR APPLICATION NUMBER: 60/092,844  
PRIOR FILING DATE: 1998-07-14  
PRIOR APPLICATION NUMBER: PCT/US99/15871  
PRIOR FILING DATE: 1999-07-13  
PRIOR APPLICATION NUMBER: 09/720383  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 1086  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-900-237-10

Query Match	100.0%	Score 5918;	DB 10;	Length 1086;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1086;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	MEASAGLVAGSHNRRLVIRRDGDPGKPREQNGOVCOICGDDVGLAPGDPVACNE	60	
DB	1	MEASAGLVAGSHNRRLVIRRDGDPGKPREQNGOVCOICGDDVGLAPGDPVACNE	60	
QY	61	CAFPYCRDCYEREREGTQNCPCQCTRYRKLKGCCORVTGDEBEDVDLDNEFNMDGDS	120	
DB	61	CAFPYCRDCYEREREGTQNCPCQCTRYRKLKGCCORVTGDEBEDVDLDNEFNMDGDS	120	
QY	121	OSVASMLYGHMSYRGDGPNGAPGAFOALNPVLLTTGQWVDDIPEQHNLLVPSFMGCG	180	
DB	121	OSVASMLYGHMSYRGDGPNGAPGAFOALNPVLLTTGQWVDDIPEQHNLLVPSFMGCG	180	

Db 121 QSVASBMLYGHNSYGRGDPNGAPQAFQOLNPVPLTLNGOMVDIPPEQHALVPSFMGG 180  
 Qy 181 GKRHPLPYADPSLPVQPSMDSPKDLAAYGYSVAMKERMENMKORBERMQTNDGG 240  
 Db 181 GKRHPLPYADPSLPVQPSMDSPKDLAAYGYSVAMKERMENMKORBERMQTNDGG 240  
 Qy 241 DGGDDADLPLMDEARQOLSRKILPSSQINPYRMIIIRLVIGFFPHYRVMHPVNDAPA 300  
 Db 241 DGGDDADLPLMDEARQOLSRKILPSSQINPYRMIIIRLVIGFFPHYRVMHPVNDAPA 300  
 Qy 301 LMLISYCEIWFAMSMILDOFPKMPPIERETYLDRLSRDKESQSQLAPIDFVSTVD 360  
 Db 301 LMLISYCEIWFAMSMILDOFPKMPPIERETYLDRLSRDKESQSQLAPIDFVSTVD 360  
 Qy 361 PLKEPPLVTNTVLSISVDYVDKVSQVSDGAAMLTFEALSETSEFAKKWVPCKRY 420  
 Db 361 PLKEPPLVTNTVLSISVDYVDKVSQVSDGAAMLTFEALSETSEFAKKWVPCKRY 420  
 Qy 421 NIEPRAPBMYFOOKIDYLDKQVAAVFRBRAMKREYEFKVRINALVAKAKVPEBGT 480  
 Db 421 NIEPRAPBMYFOOKIDYLDKQVAAVFRBRAMKREYEFKVRINALVAKAKVPEBGT 480  
 Qy 481 MDDGTPMGNVNDHPGMIQVFLGQSGGLCEGNEPLRLVYVSREKRPQYNNHKKAGAN 540  
 Db 481 MDDGTPMGNVNDHPGMIQVFLGQSGGLCEGNEPLRLVYVSREKRPQYNNHKKAGAN 540  
 Qy 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQPORFGIDR 600  
 Db 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQPORFGIDR 600  
 Qy 601 HDRYANRNVFPDINMGGLDGIQPIYVGTGCFRRQALYGYDAPTKKPPSTCNCMPK 660  
 Db 601 HDRYANRNVFPDINMGGLDGIQPIYVGTGCFRRQALYGYDAPTKKPPSTCNCMPK 660  
 Qy 661 WCECCCFGNRKQKTKTKTEKKKLLFKKEENOSPAYALGEIDEAPALENEKAGIYN 720  
 Db 661 WCECCCFGNRKQKTKTKTEKKKLLFKKEENOSPAYALGEIDEAPALENEKAGIYN 720  
 Qy 721 QOKLEKKFGQSSVFTSTLLENGTILKSPASLKEAIVHISGVEDKTDWKEIGMIY 780  
 Db 721 QOKLEKKFGQSSVFTSTLLENGTILKSPASLKEAIVHISGVEDKTDWKEIGMIY 780  
 Qy 781 GSVTEBILTPKPKHCHGMSIYCI PKVAFKGSAPNLSPRLHQLVRLMALGSIIEIFPSNH 840  
 Db 781 GSVTEBILTPKPKHCHGMSIYCI PKVAFKGSAPNLSPRLHQLVRLMALGSIIEIFPSNH 840  
 Qy 841 CPLMVGGLKFLERESYNSIYVPTSTIPLAAYCTLPALCULTGKFIPELINNVASLM 900  
 Db 841 CPLMVGGLKFLERESYNSIYVPTSTIPLAAYCTLPALCULTGKFIPELINNVASLM 900  
 Qy 901 FMSLFICIPATSILEMWGVGVIDMWRNEQFVWIGVSSHLPFAVFGGLKVLAVDTSP 960  
 Db 901 FMSLFICIPATSILEMWGVGVIDMWRNEQFVWIGVSSHLPFAVFGGLKVLAVDTSP 960  
 Qy 961 TVTSGGDBDEFSELYTFKWTLLIIPPTLLLNLFICVAVGSNAINNGESGPIFGKL 1020  
 Db 961 TVTSGGDBDEFSELYTFKWTLLIIPPTLLLNLFICVAVGSNAINNGESGPIFGKL 1020  
 Qy 1021 FPFAMVIVHLYPFLKGVGRONRPTTVIWSILLASIFSLMVRIDPLAKDGPILLES 1080  
 Db 1021 FPFAMVIVHLYPFLKGVGRONRPTTVIWSILLASIFSLMVRIDPLAKDGPILLES 1080  
 Qy 1081 CGLDCN 1086  
 Db 1081 CGLDCN 1086

RESULT 2  
 US-10-209-059-50  
 ; Sequence 50, Application US/10209059  
 ; Publication No. US2003016383A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dhugga, Kanwarpal S.

; APPLICANT: Wang, Haiyin  
 ; TITLE OF INVENTION: Maltose Cellulose Synthases and Uses  
 ; FILE REFERENCE: 0864R2  
 ; CURRENT APPLICATION NUMBER: US/10/209,059  
 ; PRIOR FILING DATE: 2002-07-31  
 ; PRIOR FILING DATE: 1998-08-17  
 ; PRIOR FILING DATE: 1999-08-06  
 ; PRIOR FILING DATE: 1999-08-06  
 ; PRIOR FILING DATE: 2000-04-14  
 ; NUMBER OF SEQ ID NOS: 52  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 50  
 ; LENGTH: 1086  
 ; TYPE: PR  
 ; ORGANISM: Zea mays  
 ; US-10-209-059-50  
 Query Match 100.0%; Score 5918; DB 12; Length 1086;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MEASAGLVAGSHNRNELVYIRRDGPGRKPPRBNQVQCQICGDDVGLAPGGDPVACNE 60  
 Db 1 MEASAGLVAGSHNRNELVYIRRDGPGRKPPRBNQVQCQICGDDVGLAPGGDPVACNE 60  
 Qy 61 CAPFCRDCYERREBGTNCPCQKTRKYLKGCORVYGBDEBDVDDLNEFNNMGDS 120  
 Db 61 CAPFCRDCYERREBGTNCPCQKTRKYLKGCORVYGBDEBDVDDLNEFNNMGDS 120  
 Qy 121 QSVASBMLYGHNSYGRGDPNGAPQAFQOLNPVPLTLNGOMVDIPPEQHALVPSFMGG 180  
 Db 121 QSVASBMLYGHNSYGRGDPNGAPQAFQOLNPVPLTLNGOMVDIPPEQHALVPSFMGG 180  
 Qy 181 GKRHPLPYADPSLPVQPSMDSPKDLAAYGYSVAMKERMENMKORBERMQTNDGG 240  
 Db 181 GKRHPLPYADPSLPVQPSMDSPKDLAAYGYSVAMKERMENMKORBERMQTNDGG 240  
 Qy 241 DGGDDADLPLMDEARQOLSRKILPSSQINPYRMIIIRLVIGFFPHYRVMHPVNDAPA 300  
 Db 241 DGGDDADLPLMDEARQOLSRKILPSSQINPYRMIIIRLVIGFFPHYRVMHPVNDAPA 300  
 Qy 301 LMLISYCEIWFAMSMILDOFPKMPPIERETYLDRLSRDKESQSQLAPIDFVSTVD 360  
 Db 301 LMLISYCEIWFAMSMILDOFPKMPPIERETYLDRLSRDKESQSQLAPIDFVSTVD 360  
 Qy 361 PLKEPPLVTNTVLSISVDYVDKVSQVSDGAAMLTFEALSETSEFAKKWVPCKRY 420  
 Db 361 PLKEPPLVTNTVLSISVDYVDKVSQVSDGAAMLTFEALSETSEFAKKWVPCKRY 420  
 Qy 421 NIEPRAPBMYFOOKIDYLDKQVAAVFRBRAMKREYEFKVRINALVAKAKVPEBGT 480  
 Db 421 NIEPRAPBMYFOOKIDYLDKQVAAVFRBRAMKREYEFKVRINALVAKAKVPEBGT 480  
 Qy 481 MDDGTPMGNVNDHPGMIQVFLGQSGGLCEGNEPLRLVYVSREKRPQYNNHKKAGAN 540  
 Db 481 MDDGTPMGNVNDHPGMIQVFLGQSGGLCEGNEPLRLVYVSREKRPQYNNHKKAGAN 540  
 Qy 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQPORFGIDR 600  
 Db 541 ALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQPORFGIDR 600  
 Qy 601 HDRYANRNVFPDINMGGLDGIQPIYVGTGCFRRQALYGYDAPTKKPPSTCNCMPK 660  
 Db 601 HDRYANRNVFPDINMGGLDGIQPIYVGTGCFRRQALYGYDAPTKKPPSTCNCMPK 660  
 Qy 661 WCECCCFGNRKQKTKTKTEKKKLLFKKEENOSPAYALGEIDEAPALENEKAGIYN 720  
 Db 661 WCECCCFGNRKQKTKTKTEKKKLLFKKEENOSPAYALGEIDEAPALENEKAGIYN 720  
 Qy 721 QOKLEKKFGQSSVFTSTLLENGTILKSPASLKEAIVHISGVEDKTDWKEIGMIY 780

```

Db 721 QOKLEKKGSSVFTSTLLENGTLKSASPASLKEAIHVISCEYEDKTMGKEIGMIY 780
Qy 781 GSVTEIDILTGFMHCHGMRISYICPKRVAFKGAPLNTSDRLHOVLKRWALGSIIEFPSNH 840
Db 781 GSVTEIDILTGFMHCHGMRISYICPKRVAFKGAPLNTSDRLHOVLKRWALGSIIEFPSNH 840
Qy 841 CPLWVGAGGGLKFLERFSYINSIYVPMWISPLLAACLTIPALCLLNGKETTPELNNVASLW 900
Db 841 CPLWVGAGGGLKFLERFSYINSIYVPMWISPLLAACLTIPALCLLNGKETTPELNNVASLW 900
Qy 901 FMSLFICIFATSILEMRSVGVIDDMWRNEQFVWIGVSSHFAVFOGLKVIAGVDTSF 960
Db 901 FMSLFICIFATSILEMRSVGVIDDMWRNEQFVWIGVSSHFAVFOGLKVIAGVDTSF 960
Qy 961 TVTSKGDDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL 1020
Db 961 TVTSKGDDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL 1020
Qy 1021 FFAPFVIVHLVPLFKGLVGRNRTPTIYVWSILASIFSLMWRIIDPLAKDGPPLLEE 1080
Db 1021 FFAPFVIVHLVPLFKGLVGRNRTPTIYVWSILASIFSLMWRIIDPLAKDGPPLLEE 1080
Qy 1081 CGLDGN 1086
Db 1081 CGLDGN 1086

```

## RESULT 3

```

US-10-160-719-18
; Sequence 18, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FaSTSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-18

```

```

Query Match 100.0%; Score 5918; DB 12; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MEASGLVAGSHNRRLVIRPDGPRKPREQNGQVCQICGDVGLAPGSDPPVAANE 60
Db 1 MEASGLVAGSHNRRLVIRPDGPRKPREQNGQVCQICGDVGLAPGSDPPVAANE 60
Qy 61 CAPPYCRDCYEXEREGTQNCPOCKTRYKLGKQCVLGDDEEDVDLNDNEFNDGHS 120
Db 61 CAPPYCRDCYEXEREGTQNCPOCKTRYKLGKQCVLGDDEEDVDLNDNEFNDGHS 120
Qy 121 QSVASMTLYGHSYGRGDPNGAPDAFQILNPVPLITNGQVDDIPEQHALVPSFMGSG 180
Db 121 QSVASMTLYGHSYGRGDPNGAPDAFQILNPVPLITNGQVDDIPEQHALVPSFMGSG 180
Qy 181 GKRIHPLVADPSLTVOPRSMDSKDIAVGYGVAMKERMNMKORBERMHQTDNDGG 240
Db 181 GKRIHPLVADPSLTVOPRSMDSKDIAVGYGVAMKERMNMKORBERMHQTDNDGG 240

```

```

Qy 241 DDGDADLPLMDARQQLSRKILPSSQINPYRMIILRLVVLGFFHTRVMMHVNDAA 300
Db 241 DDGDADLPLMDARQQLSRKILPSSQINPYRMIILRLVVLGFFHTRVMMHVNDAA 300
Qy 301 LMLISVCEIWFASWILIDPPKWFPIERETYLDRSLRFDKEQPSQLAPIDPFVSTVD 360
Db 301 LMLISVCEIWFASWILIDPPKWFPIERETYLDRSLRFDKEQPSQLAPIDPFVSTVD 360
Qy 361 PLKEPPLVNTNVLISLVDYVDKVSQVSDGAMLTBEALSETSEPAKQVPPCKRY 420
Db 361 PLKEPPLVNTNVLISLVDYVDKVSQVSDGAMLTBEALSETSEPAKQVPPCKRY 420
Qy 421 NIEPRABEYVQOKIDILDKVANKFBRERAMREIEFRVRNALVAKAQKIPESGWT 480
Db 421 NIEPRABEYVQOKIDILDKVANKFBRERAMREIEFRVRNALVAKAQKIPESGWT 480
Qy 481 MODGTPMGNVVRDHPGMIQVFLGQSGGLDCEGNELRVLVYSREKSPGVNHHKAGAMN 540
Db 481 MODGTPMGNVVRDHPGMIQVFLGQSGGLDCEGNELRVLVYSREKSPGVNHHKAGAMN 540
Qy 541 ALVVSAVLTNAPYLMLDCDHYINNSKAIKEAMCFMMDPLGKVCYVOPQRFDDIDR 600
Db 541 ALVVSAVLTNAPYLMLDCDHYINNSKAIKEAMCFMMDPLGKVCYVOPQRFDDIDR 600
Qy 601 HDRYANRVVFPDINMGDLGIGPIYVGTGCVFRQALYGYDAPKTKKPPSRTCNCMPK 660
Db 601 HDRYANRVVFPDINMGDLGIGPIYVGTGCVFRQALYGYDAPKTKKPPSRTCNCMPK 660
Qy 661 WFCFCCGFGNRKOKKTKPKTEKKKLFPKKEENOSPAYALGIDEAPAEAKAGIYN 720
Db 661 WFCFCCGFGNRKOKKTKPKTEKKKLFPKKEENOSPAYALGIDEAPAEAKAGIYN 720
Qy 721 QOKLEKKGSSVFTSTLLENGTLKSASPASLKEAIHVISCEYEDKTMGKEIGMIY 780
Db 721 QOKLEKKGSSVFTSTLLENGTLKSASPASLKEAIHVISCEYEDKTMGKEIGMIY 780
Qy 781 GSVTEIDILTGFMHCHGMRISYICPKRVAFKGAPLNTSDRLHOVLKRWALGSIIEFPSNH 840
Db 781 GSVTEIDILTGFMHCHGMRISYICPKRVAFKGAPLNTSDRLHOVLKRWALGSIIEFPSNH 840
Qy 841 CPLWVGAGGGLKFLERFSYINSIYVPMWISPLLAACLTIPALCLLNGKETTPELNNVASLW 900
Db 841 CPLWVGAGGGLKFLERFSYINSIYVPMWISPLLAACLTIPALCLLNGKETTPELNNVASLW 900
Qy 901 FMSLFICIFATSILEMRSVGVIDDMWRNEQFVWIGVSSHFAVFOGLKVIAGVDTSF 960
Db 901 FMSLFICIFATSILEMRSVGVIDDMWRNEQFVWIGVSSHFAVFOGLKVIAGVDTSF 960
Qy 961 TVTSKGDDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL 1020
Db 961 TVTSKGDDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL 1020
Qy 1021 FFAPFVIVHLVPLFKGLVGRNRTPTIYVWSILASIFSLMWRIIDPLAKDGPPLLEE 1080
Db 1021 FFAPFVIVHLVPLFKGLVGRNRTPTIYVWSILASIFSLMWRIIDPLAKDGPPLLEE 1080
Qy 1081 CGLDGN 1086
Db 1081 CGLDGN 1086

```

## RESULT 4

```

US-10-160-719-38
; Sequence 38, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: thereof

```

```

; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-160-719-38

Query Match      100.0%; Score 5918; DB 12; Length 1086;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEASAGLVAGSHNRRELVVIRRDGPGRPREQNGVCOICGDDVGLAPGDDPVAACNE 60
DB 1 MEASAGLVAGSHNRRELVVIRRDGPGRPREQNGVCOICGDDVGLAPGDDPVAACNE 60

QY 61 CAFPRCDCEYERREBGTQNCPOCKTRYKLGCCQRYVGDDEBDGVDDLDNFRMNDGHS 120
DB 61 CAFPRCDCEYERREBGTQNCPOCKTRYKLGCCQRYVGDDEBDGVDDLDNFRMNDGHS 120

QY 121 QSVASBMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWVDIPEQHALVPSFMGG 180
DB 121 QSVASBMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWVDIPEQHALVPSFMGG 180

QY 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQORRMHQTDGSGG 240
DB 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQORRMHQTDGSGG 240

QY 241 DGGDDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 300
DB 241 DGGDDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 300

QY 301 LMLISVCEIWFAMSWILDOFPKMFPIERETVLDRLSLRPDEGQSPQLAPIDFVSTVD 360
DB 301 LMLISVCEIWFAMSWILDOFPKMFPIERETVLDRLSLRPDEGQSPQLAPIDFVSTVD 360

QY 361 PLKEPPLVTTNTVLSLSDVDPVDKVSQVSDGAAMLFEALSETSEFAKXKVPCEKRY 420
DB 361 PLKEPPLVTTNTVLSLSDVDPVDKVSQVSDGAAMLFEALSETSEFAKXKVPCEKRY 420

QY 421 NIEPRAPEMYFOOKIDYLDKQVAANFVRERBRAMKREYEFKTRINALVAKAQVPEEGT 480
DB 421 NIEPRAPEMYFOOKIDYLDKQVAANFVRERBRAMKREYEFKTRINALVAKAQVPEEGT 480

QY 481 MODGTWPGNNVRDHPGMIQVFLGSGGLDCBEGNELPRLVYVSRERPGYNHHKXGAMN 540
DB 481 MODGTWPGNNVRDHPGMIQVFLGSGGLDCBEGNELPRLVYVSRERPGYNHHKXGAMN 540

QY 541 ALVRSAVLTLNAPYLTLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGDIDR 600
DB 541 ALVRSAVLTLNAPYLTLNDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRPDGDIDR 600

QY 601 HBRVYANRVVFPDINKKGLDGIQPIYVGTGCVFRQALYGDAPRTKRPSPSTCCWPK 660
DB 601 HBRVYANRVVFPDINKKGLDGIQPIYVGTGCVFRQALYGDAPRTKRPSPSTCCWPK 660

QY 661 WFCFCCCFGNRKOKTKTKTEKKLLPFKKEENOSPAYALGEIDAAAGAEKXGAYN 720
DB 661 WFCFCCCFGNRKOKTKTKTEKKLLPFKKEENOSPAYALGEIDAAAGAEKXGAYN 720

QY 721 QOKLEKFKGQSSVFTSTLLENGTLKSASPASLKEAIIHVISCGYEDTKDWKEIGMIV 780
DB 721 QOKLEKFKGQSSVFTSTLLENGTLKSASPASLKEAIIHVISCGYEDTKDWKEIGMIV 780

QY 781 GSTTEIILTFKMKHCHGWSIYCIPIKRVAFKSGAPLNLSDRLHOVLKALGSIETFFSNH 840
DB 781 GSTTEIILTFKMKHCHGWSIYCIPIKRVAFKSGAPLNLSDRLHOVLKALGSIETFFSNH 840

```

```

DB 781 GSTTEIILTFKMKHCHGWSIYCIPIKRVAFKSGAPLNLSDRLHOVLKALGSIETFFSNH 840
QY 841 CPLWYGYGGLKFLERFSYINSIYPMWISPLAAYCTLPAICLLTGKRTPELNNVSLW 900
DB 841 CPLWYGYGGLKFLERFSYINSIYPMWISPLAAYCTLPAICLLTGKRTPELNNVSLW 900
QY 901 FMSLFLICFATSIILEMWSGVGIDWNRNEQFWIYGVSSHLPAVFGGLKVIYGVDTSF 960
DB 901 FMSLFLICFATSIILEMWSGVGIDWNRNEQFWIYGVSSHLPAVFGGLKVIYGVDTSF 960
QY 961 TVTSKGDDEEFSELYPKWTTLLIPTTLILLNPIGVAGVSAINNNGESKGPPLFGKL 1020
DB 961 TVTSKGDDEEFSELYPKWTTLLIPTTLILLNPIGVAGVSAINNNGESKGPPLFGKL 1020
QY 1021 FFAFWIVLHLPPLKGLVGRQNRPTIIVWSIILASIFSLMWIRIDPFLAKDDGPLER 1080
DB 1021 FFAFWIVLHLPPLKGLVGRQNRPTIIVWSIILASIFSLMWIRIDPFLAKDDGPLER 1080
QY 1081 CGLDON 1086
DB 1081 CGLDON 1086

RESULT 5
US-09-900-237-4
; Sequence 4, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-900-237-4

Query Match      94.1%; Score 5570.5; DB 10; Length 1148;
Beet Local Similarity 94.2%; Pred. No. 0;
Matches 1027; Conservative 23; Mismatches 35; Indels 5; Gaps 3;

QY 1 MEASAGLVAGSHNRRELVVIRRDGPGRPREQNGVCOICGDDVGLAPGDDPVAACNE 60
DB 60 MEASAGLVAGSHNRRELVVIRRDGPGRPREQNGVCOICGDDVGNPGBEPVACNE 119

QY 61 CAFPRCDCEYERREBGTQNCPOCKTRYKLGCCQRYVGDDEBDGVDDLDNFRMNDGHS 120
DB 61 CAFPRCDCEYERREBGTQNCPOCKTRYKLGCCQRYVGDDEBDGVDDLDNFRMNDGHS 120

QY 120 QSVASBMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWVDIPEQHALVPSFMGG 179
DB 120 QSVASBMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWVDIPEQHALVPSFMGG 179

QY 180 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQORRMHQTDGSGG 240
DB 180 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQORRMHQTDGSGG 240

QY 240 GDDGDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 299
DB 240 GDDGDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 299

QY 300 GDDGDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 359
DB 300 GDDGDADLPLMDBARQQLSRKIPLESSQINPRMIIILVVLGFFHYRVMPVNDAA 359

QY 359 ALMLISVCEIWFAMSWILDOFPKMFPIERETVLDRLSLRPDEGQSPQLAPIDFVSTVD 359
DB 359 ALMLISVCEIWFAMSWILDOFPKMFPIERETVLDRLSLRPDEGQSPQLAPIDFVSTVD 359

```



```

Db 360 ALMLISVCEIWFAMSWILDQPKMPIERETYLRLSLRFDKEGHSQALAVDFVSTV 419
Qy 360 DPLKEPPLVNTNTLSTLSVDYPRDKVSCYSDGGAAMLTEALSTSEPKKVPCKR 419
Db 420 DPLKEPPLVNTNTLSTLSVDYPRDKVSCYSDGGAAMLTEALSTSEPKKVPCKR 479
Qy 420 YNIBRAPAWYFOOKIDYLKDKVANAFREBRAMKREBEFKVAINLVAQKVPBEGW 479
Db 480 YSLEPRAPAWYFOOKIDYLKDKVANAFREBRAMKREBEFKVAINLVAQKVPBEGW 539
Qy 480 TMOGTFMPGNVNDHPGMIQVFLQSGSGDLCEGNEPLRLVYVSREKPGYNHKKAGAM 539
Db 540 TMOGTFMPGNVNDHPGMIQVFLQSGSGDLCEGNEPLRLVYVSREKPGYNHKKAGAM 599
Qy 540 NALVRVAVLTNAAYLLNDCDHYINNSKAIKEMCFMMDLGLKCYVOPFORPGID 599
Db 600 NALVRVAVLTNAAYLLNDCDHYINNSKAIKEMCFMMDLGLKCYVOPFORPGID 659
Qy 600 RHDYRANRVYFPIINMKGLDGIQPIYVGTGCVFRROALGYDAPKPKPSTCQMP 659
Db 660 RHDYRANRVYFPIINMKGLDGIQPIYVGTGCVFRROALGYDAPKPKPSTCQMP 719
Qy 660 KMCPCCCFGNRKOK--KITKPYTEKKKLLFKKEENOSPAVALGEIDEAPGAENKA 716
Db 720 KMCPCCCFGNRKOKKITKPYTEKKKLLFKKEENOSPAVALGEIDEAPGAENKA 778
Qy 717 GIYVQOQLEKKFGGSSVFTVSTLLENGTLLKASBPASLLKEAHIVISCGIEDKTDGKEI 776
Db 779 SIYVQOQLEKKFGGSSVFTVSTLLENGTLLKASBPASLLKEAHIVISCGIEDKTDGKEI 838
Qy 777 GMIYGSYTEDILNFKHCHGMRISYCIKRVAFKGSAPLNLSDRLQOVLMALGSTEIF 836
Db 839 GMIYGSYTEDILNFKHCHGMRISYCIKRVAFKGSAPLNLSDRLQOVLMALGSTEIF 898
Qy 837 FSNHCPLMYGGGKLFELERFSYINSIVYPTWSIPLAYCTLPACLLTGKFTPELANNV 896
Db 899 FSNHCPLMYGGGKLFELERFSYINSIVYPTWSIPLAYCTLPACLLTGKFTPELANNV 958
Qy 897 ASLWFMNLFICIFATSIILEMRSGVGDIDMWRNEQFVWIGVSSHLPFAVFGGLKVIAGV 956
Db 959 ASLWFMNLFICIFATSIILEMRSGVGDIDMWRNEQFVWIGVSSHLPFAVFGGLKVIAGV 1018
Qy 957 DTSTVTSKGGDDESELYTFPKTTLLIPPTLLLNFTGVAGVGNAINNGVSGPL 1016
Db 1019 DTSTVTSKGGDDESELYTFPKTTLLIPPTLLLNFTGVAGVGNAINNGVSGPL 1078
Qy 1017 FGKLFPAFWYVHLVYPLKGLVGRONRTPRTIVVMSILLASIFSLMWRIIDPLAKDDGP 1076
Db 1079 FGKLFPAFWYVHLVYPLKGLVGRONRTPRTIVVMSILLASIFSLMWRIIDPLAKDDGP 1138
Qy 1077 LLEECGLDCN 1086
Db 1139 LLEECGLDCN 1148

```

RESULT 6  
US-10-209-059-42

```

; Sequence 42, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugge, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14

```

```

; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1059
; TYPE: PRT
; ORGANISM: Zea mays
US-10-209-059-42

Query Match          91.4%; Score 5408.5; DB 12; Length 1059;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

Qy 33 EONGVOCIGDDVGLAGDPFVACNECAFPCRCRCYEREREGTQNCQCTTRKYLK 92
Db 3 EONGVOCIGDDVGNRPDEPFVACNECAFPCRCRCYEREREGTQNCQCTTRKYLK 62
Qy 93 GCOAVTGDEBEDVDLDNEFNW-DGHDSQSVASEMLYGHMSYGRGDPGCAQAPOLN 151
Db 63 GCAVAVGDEBEDVDLDNEFNWSDGHDSYLABSLHAMSYGRGADLDGVQPHPTP 122
Qy 152 NPELITNGQVDDIPEQHALVPSFMGGGGRIRHPLPYADPSLPVOPRSMDSKDLAAYG 211
Db 123 NPELITNGQVDDIPEQHALVPSFMGGGGRIRHPLPYADPSLPVOPRSMDSKDLAAYG 182
Qy 212 YGSVAMKEBEMKOROBHQRGNDGGDDGDADLPMDERQOLSKRILPSSQINP 271
Db 183 YGSVAMKEBEMKOROBHQRGNDGGDDGDADLPMDERQOLSKRILPSSQINP 242
Qy 272 YRMIIIRLVVLCFFHYRVMHPVNDAFALMLTSVCEIWFAMSWILDQPKMPIERET 331
Db 243 YRMIIIRLVVLCFFHYRVMHPVNDAFALMLTSVCEIWFAMSWILDQPKMPIERET 302
Qy 332 YDLRLSLRFDKEGHSQALAVDFVSTVPLKEPPLVNTNTLSTLSVDYPRDKVSCYVS 391
Db 303 YDLRLSLRFDKEGHSQALAVDFVSTVPLKEPPLVNTNTLSTLSVDYPRDKVSCYVS 362
Qy 392 DDGAAMLTPALSTSEPAKQWPFCKRYNIBRAPAWYFOOKIDYLKDKVANAFREBR 451
Db 363 DDGAAMLTPALSTSEPAKQWPFCKRYNIBRAPAWYFOOKIDYLKDKVANAFREBR 422
Qy 452 AMREYEEFVRINLVAQKQVBEQWMTQDTPMPGNVVRHPGMIQVFLQSGSGDLDC 511
Db 423 AMREYEEFVRINLVAQKQVBEQWMTQDTPMPGNVVRHPGMIQVFLQSGSGDLDC 482
Qy 512 EGNELPRLVYVSREKPGYNHKKAGAMNALVRSVLTNAAYLLNDCDHYINNSKAIR 571
Db 483 EGNELPRLVYVSREKPGYNHKKAGAMNALVRSVLTNAAYLLNDCDHYINNSKAIR 542
Qy 572 EAMCFMMDPLGLKCYVOPFORFGDIIDHRYRANRVYFPIINMKGLDGIQPIYVGTG 631
Db 543 EAMCFMMDPLGLKCYVOPFORFGDIIDHRYRANRVYFPIINMKGLDGIQPIYVGTG 602
Qy 632 CVFRROALGYDAPKPKPSTCQMPKMCFCFCFNGNRKOK--KITKPYTEKKKLLP 688
Db 603 CVFRROALGYDAPKPKPSTCQMPKMCFCFCFNGNRKOKKITKPYTEKKKLLP 661
Qy 689 FKKEENOSPAVALGEIDEAPGAENKAGIVNQOQLEKKFGGSSVFTVSTLLENGTLLS 748
Db 662 FKKEENOSPAVALGEIDEAPGAENKAGIVNQOQLEKKFGGSSVFTVSTLLENGTLLS 721
Qy 749 ASPASLLKEAHIVISGYEDKTDGKEIGWYGSVTEIDILTGFKHCHGMRISYICPKRA 808
Db 722 ASPASLLKEAHIVISGYEDKTDGKEIGWYGSVTEIDILTGFKHCHGMRISYICPKRA 781
Qy 809 AFKGSAPLNLSDRLQOVLMALGSIIFPSNHCEPLVYGGGKLFELERFSYINSIYPT 868
Db 782 AFKGSAPLNLSDRLQOVLMALGSIIFPSNHCEPLVYGGGKLFELERFSYINSIYPT 841
Qy 869 SIPPLAYCTLPACLLTGKFTPELNNVASLWMSLFCIFATSIILEMRSGVGDIDMWR 928
Db 842 SIPPLAYCTLPACLLTGKFTPELNNVASLWMSLFCIFATSIILEMRSGVGDIDMWR 901
Qy 929 NEOFWVIGVSSHLPFAVFGGLKVIAGVDTSTVTSKGGDDESELYTFPKTTLLIPT 988

```

```

Db      902 NEQFVIGVSHLFAVFOGLKVIAGVDTSTTVSKGDDSESELTFFKMTLLIPT 961
Qy      989 TLLLNFIGVAVGASNAINNNGYESWGPLFGKLFPAFWIVLHLYPLKGLVGRONRTPTIV 1048
Db      962 TLLLNFIGVAVGASNAINNNGYESWGPLFGKLFPAFWIVLHLYPLKGLVGRONRTPTIV 1021
Qy      1049 IYMSILLASIFSLMWRIDPFLAKDGPILLECGLDGN 1086
Db      1022 IYMSILLASIFSLMWRIDPFLAKDGPILLECGLDGN 1059

```

## RESULT 7

```

US-10-160-719-2
; Sequence 2, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1059
; TYPE: PR
; ORGANISM: Zea mays
US-10-160-719-2

```

```

Query Match      91.4%; Score 5408.5; DB 12; Length 1059;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

```

```

Qy      33 EONGOVCOICGDDVGLAPGDDPFAVACNECAFPCRCDCYERREGONPOCKTRYKRLX 92
Db      3  QNRGOVCOICGDDVGNPDGEPFVACNECAFPCRCDCYERREGONPOCKTRYKRLX 62
Qy      93 GGORVYTGDEEDGVDDLDNEFNMV-DGHDSSVASESLYGHMSYRGCDPENGARFOALNP 151
Db      63 GCARVYTGDEEDGVDDLDNEFNMVSDGHDQYLAESMLHAMSYGRADLDGVPRPHPIR 122
Qy      152 NVPLLTNGQWVDIPPEQHALVPSFMGCGGKRIHPLPYADPSLIPVQPRSMDSKDLAAYG 211
Db      123 NVPLLTNGQWVDIPPEQHALVPSFMGCGGKRIHPLPYADPSLIPVQPRSMDSKDLAAYG 182
Qy      212 YGSVAMKEMKMKQKQERHQGNQGGDDGDALPLMDERQOLSKTILPSSQINP 271
Db      183 YGSVAMKEMKMKQKQERHQGNQGGDDGDALPLMDERQOLSKTILPSSQINP 242
Qy      272 YKMTIIRLVLGFFPHYVMHVNDAFALMLISVCEIWFAMSWLLDOPPKKPIERET 331
Db      243 YKMTIIRLVLCFFPHYVMHVNDAFALMLISVCEIWFAMSWLLDOPPKKPIERET 302
Qy      332 YDRLSLRFDKEGQPSQLAPIDFVSTVDPLEKEPILVTNTVLISLVDPVDKVSQYVS 391
Db      303 YDRLSLRFDKEGQPSQLAPIDFVSTVDPLEKEPILVTNTVLISLVDPVDKVSQYVS 362
Qy      392 DDGAAMLTPEALSETSEPAKWPFCCKRYNIERAEWYQOQIDYLKQVAFNFRERR 451
Db      363 DDGAAMLTPEALSETSEPAKWPFCCKRYNIERAEWYQOQIDYLKQVAFNFRERR 422
Qy      452 AMREVEEFKVRINALVAKQKVPBEQMTQDSTPMGNVVRHDPGMIQVFLQSGSLDLC 511
Db      423 AMREVEEFKVRINALVAKQKVPBEQMTQDSTPMGNVVRHDPGMIQVFLQSGSGHIV 482

```

```

Qy      512 EGNELPLVYVSREKRCGYNHHKKAAMNLYRVAULTNAPYLNLDDCHYINNKAIX 571
Db      483 EGNELPLVYVSREKRCGYNHHKKAAMNLYRVAULTNAPYLNLDDCHYINNKAIX 542
Qy      572 EAMCFMMDPLGKKYCVYQFPQFEDGIDRHDYRANRVVFPDINMKGLDGIQGIYVGTG 631
Db      543 EAMCFMMDPLGKKYCVYQFPQFEDGIDRHDYRANRVVFPDINMKGLDGIQGIYVGTG 602
Qy      632 CVPRQALYGDAPKTKKPSRTCCNCPKWCFCFCCCFQNRKOK--KTKRKTEKKLLF 688
Db      603 CVPRQALYGDAPKTKKPSRTCCNCPKWCICCCCFQNRKTKTKTKSKPFKIKL- 661
Qy      689 FKKEENQSPALAGEIDEAAGANEKAGIYNQKLEKKFGQSSVPTSTILENGTILKS 748
Db      662 FKKEENQSPALAGEIDEAAGANEKASYNQKLEKKFGQSSVFASTILENGTILKS 721
Qy      749 ASPASILKEAHLVHISCGVEDKTDGKEIGMTYGSVTEIDILTFQNGHGMRSIYCIPIKV 808
Db      722 ASPASILKEAHLVHISCGVEDKTDGKEIGMTYGSVTEIDILTFQNGHGMRSIYCIPIKV 781
Qy      809 AFKGSAPLNLSDRLHQLRMALGSIELFFSNHCEPLMTYGGGLKFLERFSYINSIYVPT 868
Db      782 AFKGSAPLNLSDRLHQLRMALGSIELFFSNHCEPLMTYGGGLKFLERFSYINSIYVPT 841
Qy      869 SIPLAYCTLPALCILTCKFTTPELNNVNASLWMSLFCIFATSTILEMRAGVGIDDMWR 928
Db      842 SIPLAYCTLPALCILTCKFTTPELNNVNASLWMSLFCIFATSTILEMRAGVGIDDMWR 901
Qy      929 NEQFVIGVSHLFAVFOGLKVIAGVDTSTTVSKGDDSESELTFFKMTLLIPT 988
Db      902 NEQFVIGVSHLFAVFOGLKVIAGVDTSTTVSKGDDSESELTFFKMTLLIPT 961
Qy      989 TLLLNFIGVAVGASNAINNNGYESWGPLFGKLFPAFWIVLHLYPLKGLVGRONRTPTIV 1048
Db      962 TLLLNFIGVAVGASNAINNNGYESWGPLFGKLFPAFWIVLHLYPLKGLVGRONRTPTIV 1021
Qy      1049 IYMSILLASIFSLMWRIDPFLAKDGPILLECGLDGN 1086
Db      1022 IYMSILLASIFSLMWRIDPFLAKDGPILLECGLDGN 1059

```

## RESULT 8

```

US-10-209-059-18
; Sequence 18, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugra, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1094
; TYPE: PR
; ORGANISM: Zea mays
US-10-209-059-18

```

```

Query Match      88.3%; Score 5228; DB 12; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

```

```

Qy      1 MEASAGLVAGSHNRRELTVIRRD---GDPGRPREQNGVQICGDDVGLAPGDDPFAV 57

```

```

DB 1 MEASAGLVAGSHNNELVIRDRRESGAAGGAARAEAP-CQICGDEYGVGFODEPFA 59
QY 58 CNECAFVPCRDCEYERREBGTQNCPOCKTRYRLKGCORVGTDEEBDGVDDLDNEFN-WD 116
DB 60 CNECAFVPCRACYEYERREBGSQACPOCKTRYRLKGCORVAGDEEBDGVDDLDNEFN 119
QY 117 G----HDSQSAEBSMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWMDIPPEQAL 172
DB 120 GAHEDDPQYVAESMLRAQMSYGRGD--AHGFSVPVNPVLLTNGQWMDIPPEQAL 176
QY 173 VPSFM---GGGGRJHPLPYADPSLPVOPRSMDSKDLAAYGYGSVAMKEREMKOKOE 229
DB 177 VPSFMGGGGGGRJHPLPYADPSLPVOPRSMDSKDLAAYGYGSVAMKEREMKOKOE 236
QY 230 RMHQTGNDGGD-DGDDADPLMDBARQOLSRKILPSSQINPYRMIIIRLVVGFPH 288
DB 237 RLQHRSGGGGMDDDADPLMDBARQOLSRKILPSSQINPYRMIIIRLVVGFPH 296
QY 289 YRVHAPVNDAPALMLISYCEIWFAMSWILDOPFKMPLIERETYLDRSLRFDKGGQPSQ 348
DB 297 YRVHAPVNDAPALMLISYCEIWFAMSWILDOPFKMPLIERETYLDRSLRFDKGGQPSQ 356
QY 349 LAPIDFVSTVDPLEKPLVTTNTVLTSLISVDYVVKVSCYVSDGAAMLTFEALSTSE 408
DB 357 LAPIDFVSTVDPLEKPLVTTNTVLTSLISVDYVVKVSCYVSDGAAMLTFEALSTSE 416
QY 409 FAKKVPFCRKYNIIEPRAPWYFOOKIDYLDKQVAAFVRRERAMKREYEFKVRINLV 468
DB 417 FAKKVPFCRKYNIIEPRAPWYFOOKIDYLDKQVAAFVRRERAMKREYEFKVRINLV 476
QY 469 AKAQVPEBGTWODGTWPGANNVRDHPGMIQVFLGSGGGLDCBGNELPRLVYSREKRP 528
DB 477 AKAQVPEBGTWODGTWPGANNVRDHPGMIQVFLGSGGGLDCBGNELPRLVYSREKRP 536
QY 529 GYNHKKAGANMALVRSAVLTNAAYLLNLDCHYINNSKAIKEAMCFMDPLVKKVCY 588
DB 537 GYNHKKAGANMALVRSAVLTNAAYLLNLDCHYINNSKAIKEAMCFMDPLVKKVCY 596
QY 589 VOPFORPQIDRHDYRANRVVFDINMKGLDGIQPIYVGVCFRRQALGYAPATK 648
DB 597 VOPFORPQIDRHDYRANRVVFDINMKGLDGIQPIYVGVCFRRQALGYAPATK 656
QY 649 KPRSTCNMPCMCSCCCCFGNRKOKTKPTKEKKKLFPKXENOSPAVALGEIDEA 708
DB 657 KPRSTCNMPCMCSCCCCFGNRKOKTKPTKEKKKLFPKXENOSPAVALGEIDEA 716
QY 709 PGEAEKAGIYVQOULEKKFGSSVFTVSTLLENGTLKASAPASLKAIVISGTYED 768
DB 717 PGEAEKAGIYVQOULEKKFGSSVFTVSTLLENGTLKASAPASLKAIVISGTYED 776
QY 769 KTDWKEIGMIGYSTEDILTFGKMHGMSIYCI PKRVAFKGSAPLNLSDRLQVLRM 828
DB 777 KTDWKEIGMIGYSTEDILTFGKMHGMSIYCI PKRVAFKGSAPLNLSDRLQVLRM 836
QY 829 ALGSIIEFSSNHCPLMYGGGLKFLERESYINSIYVWTSIPLAYCTLPACILTLGKF 888
DB 837 ALGSIIEFSSNHCPLMYGGGLKFLERESYINSIYVWTSIPLAYCTLPACILTLGKF 896
QY 889 ITBELNNTASLWMSLFCIFATSLLENKMSGVGIDMMWRNBOFVWIGVSHLFAVROG 948
DB 897 ITBELNNTASLWMSLFCIFATSLLENKMSGVGIDMMWRNBOFVWIGVSHLFAVROG 956
QY 949 LKVLAVGDTSTVYVSKGDEEPELTYFKMTTLLIPTTLLNFIIGVAVGNAIINN 1008
DB 957 LKVLAVGDTSTVYVSKGDEEPELTYFKMTTLLIPTTLLNFIIGVAVGNAIINN 1016
QY 1009 GYESNGPLFGKLPFAFWIVLHLYPFLKGLVGRONTPTIIVMSILLASIFSLVWRIDP 1068
DB 1017 GYESNGPLFGKLPFAFWIVLHLYPFLKGLVGRONTPTIIVMSILLASIFSLVWRIDP 1076
QY 1069 FLAKDGPILLECGLDNC 1086

```

```

DB 1077 FLAKSNGPLLECGLDNC 1094
RESULT 9
US-10-160-719-26
; Sequence 26, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthesases and Uses
; TITLE OF INVENTION: Theoreof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-26
Query Match 88.3%; Score 5228; DB 12; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;
QY 1 MEASAGLVAGSHNNELVIRDR--GDPGKPREBQNGVQCIQCGDVGLAGDGPDEPFA 57
DB 1 MEASAGLVAGSHNNELVIRDRRESGAAGGAARAEAP-CQICGDEYGVGFODEPFA 59
QY 58 CNECAFVPCRDCEYERREBGTQNCPOCKTRYRLKGCORVGTDEEBDGVDDLDNEFN-WD 116
DB 60 CNECAFVPCRACYEYERREBGSQACPOCKTRYRLKGCORVAGDEEBDGVDDLDNEFN 119
QY 117 G----HDSQSAEBSMLYGHMSYGRGDPNGAPQAFQALPNVPLLTNGQWMDIPPEQAL 172
DB 120 GAHEDDPQYVAESMLRAQMSYGRGD--AHGFSVPVNPVLLTNGQWMDIPPEQAL 176
QY 173 VPSFM---GGGGRJHPLPYADPSLPVOPRSMDSKDLAAYGYGSVAMKEREMKOKOE 229
DB 177 VPSFMGGGGGGRJHPLPYADPSLPVOPRSMDSKDLAAYGYGSVAMKEREMKOKOE 236
QY 230 RMHQTGNDGGD-DGDDADPLMDBARQOLSRKILPSSQINPYRMIIIRLVVGFPH 288
DB 237 RLQHRSGGGGMDDDADPLMDBARQOLSRKILPSSQINPYRMIIIRLVVGFPH 296
QY 289 YRVHAPVNDAPALMLISYCEIWFAMSWILDOPFKMPLIERETYLDRSLRFDKGGQPSQ 348
DB 297 YRVHAPVNDAPALMLISYCEIWFAMSWILDOPFKMPLIERETYLDRSLRFDKGGQPSQ 356
QY 349 LAPIDFVSTVDPLEKPLVTTNTVLTSLISVDYVVKVSCYVSDGAAMLTFEALSTSE 408
DB 357 LAPIDFVSTVDPLEKPLVTTNTVLTSLISVDYVVKVSCYVSDGAAMLTFEALSTSE 416
QY 409 FAKKVPFCRKYNIIEPRAPWYFOOKIDYLDKQVAAFVRRERAMKREYEFKVRINLV 468
DB 417 FAKKVPFCRKYNIIEPRAPWYFOOKIDYLDKQVAAFVRRERAMKREYEFKVRINLV 476
QY 469 AKAQVPEBGTWODGTWPGANNVRDHPGMIQVFLGSGGGLDCBGNELPRLVYSREKRP 528
DB 477 AKAQVPEBGTWODGTWPGANNVRDHPGMIQVFLGSGGGLDCBGNELPRLVYSREKRP 536
QY 529 GYNHKKAGANMALVRSAVLTNAAYLLNLDCHYINNSKAIKEAMCFMDPLVKKVCY 588
DB 537 GYNHKKAGANMALVRSAVLTNAAYLLNLDCHYINNSKAIKEAMCFMDPLVKKVCY 596

```

```

QY 589 VOEPQRFDDIDRHRANRVVFPDINMKGLDGIQPIYVGCVERROALGYDAPKTK 648
DB 597 VOEPQRFDDIDRHRANRVVFPDINMKGLDGIQPIYVGCVERROALGYDAPKTK 656
QY 649 KPESRTCNCPKWCFCFCCCGFGRNKKTKTKTEKKKLLPFKKEENOSPAYALGEIDEAA 708
DB 657 KPESRTCNCPKWCFCFCCCGFGRNKKTKTKTEKKKLLPFKKEENOSPAYALGEIDEAA 716
QY 709 PGANEKAGIVNOQKLEKFKFGSSVPTSTLLENGTLLKASAPASLKEAIIHIVISGCEYD 768
DB 717 PGADIEKAGIVNOQKLEKFKFGSSVPTSTLLENGTLLKASAPASLKEAIIHIVISGCEYD 776
QY 769 KTDWKEIGWISVSTEDILTGFKMGCHGMRSTYCI PKRYA FKGSA PLINISDLHQVLRW 828
DB 777 KTDWKEIGWISVSTEDILTGFKMGCHGMRSTYCI PKRYA FKGSA PLINISDLHQVLRW 836
QY 829 ALGSIIEIFSNHCPMLWYGCGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLLTGKF 888
DB 837 ALGSIIEIFSNHCPMLWYGCGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLLTGKF 896
QY 889 ITBELNNVASLWFMSLFICIFATSILEMNRSGVGDIDMWRNEQFWIIGVSHLFAVFOG 948
DB 897 ITBELNNVASLWFMSLFICIFATSILEMNRSGVGDIDMWRNEQFWIIGVSHLFAVFOG 956
QY 949 LKAVIAGVDTSTFTVTSKGGDEESELTYTKWTTLLIPPTLLLLNFIVGAVGSAINN 1008
DB 957 LKAVIAGVDTSTFTVTSKGGDEESELTYTKWTTLLIPPTLLLLNFIVGAVGSAINN 1016
QY 1009 GYESWGPLEFGKLFPAFWIVLHLPFLKGLVGRONRPTIIVWSILLASIFSLWVRIDP 1068
DB 1017 GYESWGPLEFGKLFPAFWIVLHLPFLKGLVGRONRPTIIVWSILLASIFSLWVRIDP 1076
QY 1069 FLAKDGPILLECGLCDN 1086
DB 1077 FLAKSNGPLLECGLCDN 1094

RESULT 10
US-10-160-719-46
; Sequence 46, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kamarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160. 719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-160-719-46

Query Match 88.3%; Score 5228; DB 12; Length 1094;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 36; Mismatches 70; Indels 16; Gaps 7;
QY 1 MEASAGLVAGSHNNELNVLIRRD---GDPGKPREPREGQVQICGDDVGLAPGDBFVA 57
DB 1 MEASAGLVAGSHNNELNVLIRRDREGAGAGGARRAEAP-CQCGDEVGVGFGGEPFVA 59
QY 58 CNECAFVPCRDYEREREGONCPQCKTRYKRLKGGQRTYGDDEENGVDLNDNERN-WD 116

```

```

DB 60 CNECAFVPCRDYEREREGONCPQCKTRYKRLKGGQRTYGDDEENGVDLNDNERN-WD 119
QY 117 G----HDSQSVAESNLVGHMSTYGRGDPNGAPQALPNPPLTNGQVMDIPEEQNAL 172
DB 120 GAHEHDDPQYVAESNLVGHMSTYGRGDPNGAPQALPNPPLTNGQVMDIPEEQNAL 176
QY 173 VPSFM---GGGGRKHPLPYADPSLPVQPRSMDEPKDLAAAGVGVANKEHEMKNKOE 229
DB 177 VPSFM---GGGGRKHPLPYADPSLPVQPRSMDEPKDLAAAGVGVANKEHEMKNKOE 236
QY 230 RMHQTGNDGGD-DGDDADLPLMDARQOLSRKILPLPSQINPYEMIIIRLVVLFQFFH 288
DB 237 RLGHRSEGGGWDGDDADLPLMDARQOLSRKILPLPSQINPYEMIIIRLVVLFQFFH 296
QY 289 YVWHPVDAFALMLISYCEIWFPMWSILDOFPWFPIRETTYIDRLSLFDEKQPSQ 348
DB 297 YVWHPVDAFALMLISYCEIWFPMWSILDOFPWFPIRETTYIDRLSLFDEKQPSQ 356
QY 349 LAPIDFVSTVPLKEPPLVTNTVLSISVDYPAVDKXSCVSDGAMLPFEALSETSE 408
DB 357 LAPIDFVSTVPLKEPPLVTNTVLSISVDYPAVDKXSCVSDGAMLPFEALSETSE 416
QY 409 FAKKWPFCRKNYIEPRAPWYFOQKIDYLDKQVANAIFERRAKREYEERKYNALV 468
DB 417 FAKKWPFCRKNYIEPRAPWYFOQKIDYLDKQVANAIFERRAKREYEERKYNALV 476
QY 469 AQAQYVEEGWMDQCTPAGNNDPDMQTOVPLTQSGGDLDCBNELPRLVYVERKRP 528
DB 477 AQAQYVEEGWMDQCTPAGNNDPDMQTOVPLTQSGGDLDCBNELPRLVYVERKRP 536
QY 529 GYNHKKAGAMNALRVSAVLTNAPYLLNDDHYNINSKAIKEMCFMMDPLDKKVCY 588
DB 537 GYNHKKAGAMNALRVSAVLTNAPYLLNDDHYNINSKAIKEMCFMMDPLDKKVCY 596
QY 589 VOEPQRFDDIDRHRANRVVFPDINMKGLDGIQPIYVGCVERROALGYDAPKTK 648
DB 597 VOEPQRFDDIDRHRANRVVFPDINMKGLDGIQPIYVGCVERROALGYDAPKTK 656
QY 649 KPESRTCNCPKWCFCFCCCGFGRNKKTKTKTEKKKLLPFKKEENOSPAYALGEIDEAA 708
DB 657 KPESRTCNCPKWCFCFCCCGFGRNKKTKTKTEKKKLLPFKKEENOSPAYALGEIDEAA 716
QY 709 PGANEKAGIVNOQKLEKFKFGSSVPTSTLLENGTLLKASAPASLKEAIIHIVISGCEYD 768
DB 717 PGADIEKAGIVNOQKLEKFKFGSSVPTSTLLENGTLLKASAPASLKEAIIHIVISGCEYD 776
QY 769 KTDWKEIGWISVSTEDILTGFKMGCHGMRSTYCI PKRYA FKGSA PLINISDLHQVLRW 828
DB 777 KTDWKEIGWISVSTEDILTGFKMGCHGMRSTYCI PKRYA FKGSA PLINISDLHQVLRW 836
QY 829 ALGSIIEIFSNHCPMLWYGCGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLLTGKF 888
DB 837 ALGSIIEIFSNHCPMLWYGCGGLKFLERFSYINSIYVPMWTSIPLAYCTLPALICLLTGKF 896
QY 889 ITBELNNVASLWFMSLFICIFATSILEMNRSGVGDIDMWRNEQFWIIGVSHLFAVFOG 948
DB 897 ITBELNNVASLWFMSLFICIFATSILEMNRSGVGDIDMWRNEQFWIIGVSHLFAVFOG 956
QY 949 LKAVIAGVDTSTFTVTSKGGDEESELTYTKWTTLLIPPTLLLLNFIVGAVGSAINN 1008
DB 957 LKAVIAGVDTSTFTVTSKGGDEESELTYTKWTTLLIPPTLLLLNFIVGAVGSAINN 1016
QY 1009 GYESWGPLEFGKLFPAFWIVLHLPFLKGLVGRONRPTIIVWSILLASIFSLWVRIDP 1068
DB 1017 GYESWGPLEFGKLFPAFWIVLHLPFLKGLVGRONRPTIIVWSILLASIFSLWVRIDP 1076
QY 1069 FLAKDGPILLECGLCDN 1086
DB 1077 FLAKSNGPLLECGLCDN 1094

RESULT 11
US-09-900-237-8

```

```

; Sequence 8, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BBI170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 1165
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-900-237-8

```

```

Query Match      88.3%; Score 5228; DB 10; Length 1165;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

```

```

QY 1 MEASAGLVAGSHNRNELVIRRD---GDPGPKPREQNGVCQICGDDVGLAPGDDPFA 57
DB 72 MEASAGLVAGSHNRNELVIRRDREGAAGGAARAEAP-CQICDBEVGVGDFDEPFA 130
QY 58 CNECAFVPCRCCEYERREGTONCPQCKTRYKRLKGCORVTGBEBEDGVDDIDNEFN-WD 116
DB 131 CNECAFVPCRCCEYERREGTONCPQCKTRYKRLKGCORVTGBEBEDGVDDIDNEFN-WD 190
QY 117 G-----HDSQVASEMLYGHMSYGRGDPNGAPQAFQALNPVPLTLNGQVVDIIPRQHAL 172
DB 191 GAHEDBDPQYVASEMLYGHMSYGRGDPNGAPQAFQALNPVPLTLNGQVVDIIPRQHAL 247
QY 173 VPSFM---GGGGRTHPLPYADPSLPVQSRMSPSGLAAGYGVAKEMENMKOROE 229
DB 248 VPSYMSGGGGGGRTHPLPYADPSLPVQSRMSPSGLAAGYGVAKEMENMKOROE 307
QY 220 RHQGTGNDGGD-DGDDADLPLMDEARQOLSRKIPSSQINPYMIIIRLVIGFFPH 288
DB 308 RHQGTGNDGGD-DGDDADLPLMDEARQOLSRKIPSSQINPYMIIIRLVIGFFPH 367
QY 289 YRVHPRVNDAPALMLISVLCIWMFAMSWILDQFPKFPPIEREYLDRLSLRDKESQPSQ 348
DB 368 YRVHPRVNDAPALMLISVLCIWMFAMSWILDQFPKFPPIEREYLDRLSLRDKESQPSQ 427
QY 349 LAPIDFVSTVPLPKPLVTTNTVLSISVDYPVQKVSQVSDGAMLTREALSETSE 408
DB 428 LAPIDFVSTVPLPKPLVTTNTVLSISVDYPVQKVSQVSDGAMLTREALSETSE 487
QY 409 FAKKVVPRPKRYNIBPRAEWYFOOKIDYLKDKVANFPRERAKREYERKATINLY 468
DB 488 FAKKVVPRPKRYNIBPRAEWYFOOKIDYLKDKVANFPRERAKREYERKATINLY 547
QY 469 AKAQVPEEGMTMODGTPWPGANNVDPHGMIOVPLQSGSGGLCEGNEPLRLVYVREKRP 528
DB 548 AKAQVPEEGMTMODGTPWPGANNVDPHGMIOVPLQSGSGGLCEGNEPLRLVYVREKRP 607
QY 529 GYNHKKAGAMNALVRSVAVLTNAPYLLNLDGDHYIINNSKAIKEAMCFMMDPLAKRYCY 588
DB 608 GYNHKKAGAMNALVRSVAVLTNAPYLLNLDGDHYIINNSKAIKEAMCFMMDPLAKRYCY 667
QY 589 VQPPORFOSIDHDVYANRVVFPDINMKGLDGIQIPVGVGCVPRPALAGYAPATK 648
DB 668 VQPPORFOSIDHDVYANRVVFPDINMKGLDGIQIPVGVGCVPRPALAGYAPATK 727
QY 649 KPSRTQNCWPKMCECCCFGNRKQKTKTKPTEKKKLFPKEENQSPAYALGIDEA 708
DB 728 KPSRTQNCWPKMCECCCFGNRKQKTKTKPTEKKKLFPKEENQSPAYALGIDEA 787

```

```

QY 709 PGAENEKAGIVNOCKLEKKFGQSSVFVNTSTLLENGCTLKSASPALKEAHVISCYED 768
DB 768 PGADIEKAGIVNOCKLEKKFGQSSVFVNTSTLLENGCTLKSASPALKEAHVISCYED 847
QY 769 KTDWKEIGWTVGVTEDILNGFKMHCHGMSIYICPRVAFKGSAPLNTSDRLHQVLRW 828
DB 848 KTDWKEIGWTVGVTEDILNGFKMHCHGMSIYICPRVAFKGSAPLNTSDRLHQVLRW 907
QY 829 ALGSTEIFPSNHCPLWYGGGLKFLERFSSYINSIVYPMTSIPLIAYCTLPALICLLTQKF 888
DB 908 ALGSTEIFPSNHCPLWYGGGLKFLERFSSYINSIVYPMTSIPLIAYCTLPALICLLTQKF 967
QY 889 ITPELNNVASLWPMELFCIPATSLERMSGVGIDMWRBQFVAVIGVSHLPAVQ 948
DB 968 ITPELNNVASLWPMELFCIPATSLERMSGVGIDMWRBQFVAVIGVSHLPAVQ 1027
QY 949 LLKVIAGVDTFTVYSKGGDDEEFSELVTFPMKTTLLIPTLLLNFIQVAVASNAINN 1008
DB 1028 LLKVIAGVDTFTVYSKGGDDEEFSELVTFPMKTTLLIPTLLLNFIQVAVASNAINN 1087
QY 1009 GYESNGPLFGKLFPAFWYVHLYPPLKGLVGRONTPTIVIVMSILLASIFSLWVRIDP 1068
DB 1088 GYESNGPLFGKLFPAFWYVHLYPPLKGLVGRONTPTIVIVMSILLASIFSLWVRIDP 1147
QY 1069 FLAKDGPPLBERCGIDCN 1086
DB 1148 FLAKDGPPLBERCGIDCN 1165

```

## RESULT 12

```

US-10-229-193-8
; Sequence 8, Application US/10229193
; Publication No. US20030126643A1
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio; Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; TITLE OF INVENTION: Manipulation of Cellulose and/or Beta-1,4-Glucan
; FILE REFERENCE: 96-98A
; CURRENT APPLICATION NUMBER: US/10/229,193
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/4221,013
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-10-229-193-8

```

```

Query Match      71.8%; Score 4250.5; DB 15; Length 1084;
Best Local Similarity 70.8%; Pred. No. 0;
Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;

```

```

QY 1 MEASAGLVAGSHNRNELVIRRDGDPKPREQNGVCQICGDDVGLAPGDDPFA 60
DB 1 MNTGRLTAGSHNRNELVIRRDGDPKPREQNGVCQICGDDVGLAPGDDPFA 60
QY 61 CAPFVPCRCCEYERREGTONCPQCKTRYKRLKGCORVTGBEBEDGVDDIDNEFN-WD 119
DB 61 CAPFVPCRCCEYERREGTONCPQCKTRYKRLKGCORVTGBEBEDGVDDIDNEFN-WD 119
QY 120 SOSVAESMLYGHMSYGRGDPNGAPQAFQALNPVPLTLNGQVVDIIPRQHALVPSFWG 179
DB 120 PEHAEAALSSRLNTGRGLDSAPP-----GQIPLLTYCEDADDMVSDRALTVPSGTG 174
QY 180 GKKIHPLPYADPSLPVQSRMSPSGLAAGYGVAKEMENMKOROE-ERM---HQT 234

```

Db 175 YGNRYVPAFTDSSAPPOARKSVPOKIDIAEYGVSVAMKDNEMKKEGCELTQYIKHEG 234  
QY 225 GNDGGG---DDGDDADLPLMDAARQOLSRKIPLPSSQINPYRMIIIRLVYLGFFHYR 290  
Db 225 GNNNGSGDDDLDDPDMPMWDEGQPLSRKLPINSSRNINPYRMILORLILGIFHYR 294  
QY 291 VMHPVNDAPALMLISVIGEIFPAMGWLDPFKWPIRETYLDELISLFDKESQOLA 350  
Db 295 ILHPVNDAYGLMLTVCIEIMFAVSMIIDQPFKWPPIRETYLDELISLRYEKEGPRSGLA 354  
QY 351 PIDEFVSTVDPKPEPLVTNTVTLISVDPVDKSCVSDGGAAMLFEALSETSEPA 410  
Db 355 PVDVFVSTVDPKPEPLVTANTVLSLAVDIPVDKACVSNNGAAMLTPELSTADPA 414  
QY 411 KKMVFCKRYNIEPPAPEMWFOOKIDYLDKQVAAAFVRRERAMKEEYEFKRYNALYAK 470  
Db 415 TKWVFCKKFNIEPPAPEMWFSQKMDYLNKQVHAPFVRRERAMKDYEEFKYNALYAT 474  
QY 471 AOKVBEQWTMODGTPWPGNNVVDHPGMTQVFLGOSGGLDCGNEPLRLVYISREKRGY 530  
Db 475 AOKVBEQWTMODGTPWPGNNVVDHPGMTQVFLGHSVVDITDGNELPLVYISREKRPQF 534  
QY 531 NHHKKAGAMNALVRSALVLTNAPLYLNLCDHYINNSKAIKEAMGPMMDPLLGKVCYVQ 590  
Db 535 DHKKAAGAMNSLIRVSAVLNAPYLLNVDCDHYINNSKAIRESMCMMDPQSGKVCYVQ 594  
QY 591 PFOREFDIDRHRVYANRVVFFDINMGDLGIGPIYVGTGCVFRROALYGYDAPKTKP 650  
Db 595 FQORFDGIDRHRYSRNVVFFDINMGDLGIGPIYVGTGCVFRROALYGYDAPKTKP 654  
QY 651 PRTONCWPKRCFCCCFGNRKOKKT--TKPTKCKLLFFKKENQSAVALGELDEA- 707  
Db 655 POKTNCWPKRC--CLCC--GLRKSSTKATDKKTNK-----ETSKOIHLELVNDEGV 704  
QY 708 -APGAENKAGIYNOQKLEKKFGQSVFVTSLTLENGTLKASPASLKEALHYISGY 766  
Db 705 IYVSVNVERKRESEATQKLEKKFGQSPVAVASAVLQNGVPRNASPCLREAIQVYISGY 764  
QY 767 EDKTDWKEIGMIGSVTEEDILTFMCHGKMSIICIPKRYAFKGSAPLNDRLHQL 826  
Db 765 QDKTEWKEIGMIGSVTEEDILTFMCHGKMSIICIPKRYAFKGSAPLNDRLHQL 824  
QY 827 RNALGSEIFFSNHCPLWYGGGLKELERFSYINSIYVWTSIPLLAYCTLPALCLLTG 886  
Db 825 RNALGSEIFFSNHCPLWYGGGLKELERFSYINSIYVWTSIPLLAYCTLPALCLLTG 884  
QY 887 KRTTPELANVASIMFSLPCTFATSIEMRMSGVGIDDMWRNEQFWIGVSSHFAVF 946  
Db 885 KEIYPEISNYAGILFVLMFISIAVTGILEMOWGVGIDDMWRNEQFWIGVSSHFAVF 944  
QY 947 OGILKVIAGVDTSFVTSKGDDESELYTFKMTLLIPTTLLILNFIQVAGVSNAI 1006  
Db 945 OGILKVIAGVNTNFVTSKADGASELYTFKMTLLIPTTLLILNFIQVAGVSDAI 1004  
QY 1007 NNGYESWGLPFGKLFPAFVYVHLVPLKGLVGRQNRTPYIVWISILASIFSLMWRI 1066  
Db 1005 SNGYOSWGLPFGKLFPAFVYVHLVPLKGLVGRQNRTPYIVWISILASIFSLMWRI 1064  
QY 1067 DPEFLADGDLLECGLDLC 1085  
Db 1065 NPEFVAK--GGPVLIEICGLNC 1082

FILE REFERENCE: 0864R2  
CURRENT APPLICATION NUMBER: US/10/209,059  
CURRENT FILING DATE: 2002-07-31  
PRIOR APPLICATION NUMBER: 60/096,822  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 09/371,383  
PRIOR FILING DATE: 1999-08-06  
PRIOR APPLICATION NUMBER: 09/550,483  
PRIOR FILING DATE: 2000-04-14  
NUMBER OF SEQ ID NOS: 52  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1075  
TYPE: PRT  
ORGANISM: Zea mays  
US-10-209-059-2  
Query Match 67.4%; Score 3991; DB 12; Length 1075;  
Beat Local Similarity 66.5%; Pred. No. 0;  
Matches 739; Conservative 143; Mismatches 166; Indels 64; Gaps 15;  
QY 1 MEASAGLVAGSHNNELVIRDD-PG-PKPREQNGVQICGDDVGLAPGDPPVAC 58  
Db 1 MAAMGMVAGSHNNEFVIMRHDDVGSAPKTSANGVQCIGDSVGSATGDFVAC 60  
QY 59 NECAFVPCRCYCYERBRGTQNCPOCKTRYRLKGCQRYVTDDEEDVDLDNEFPMDG 118  
Db 61 NECAFVPCRCYCYERBRGTQNCPOCKTRYRKQSGSPVHDEDEEDVDLDNEFNY--- 117  
QY 119 DQSAVASEMLVGHSHYSG-----GDPNGAPQAFQALPN--VPLLTNGQWVD----DIP 166  
Db 118 -----KQSGKGPWQLOGDDADLSSARHEPHHRIIRLISGQISGEIPAS 165  
QY 167 PEOHALVSEFNGGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGVSAKEMENKQ 226  
Db 166 PDHRSI-----RPTSSVDSVPVRLVDPKDLNSGLNSVDMKEREVSRY 215  
QY 227 ROER-MHGTG-----DGGDDGDDADLPLMDAARQOLSRKIPLPSSQINPYRMI 275  
Db 216 KQDKMMQVTKRYPEARGDMEGTSGNGE--DMQWVDDARPLSRIVPISNQNLVRYV 273  
QY 276 IIRLVYLGFFHYRVMHPVNDAPALMLISVIGEIFPAMGWLDPFKWPIRETYLDR 335  
Db 274 IIRLVYLGFFHYRVMHPVNDAPALMLISVIGEIFPAMGWLDPFKWPIRETYLDR 333  
QY 336 LSLFDEKQSPQALPIDFVSTVDPKPEPLVTNTVTLISVDPVDKSCVSDGGA 395  
Db 334 LAKRDREGEQSOLAPIDVFTVDPKPEPLVTANTVLSLAVDIPVDKSCVSDGGS 393  
QY 396 AMLTFEALSETSEPAKKVVPCKRYNIEPPAPEMWFOOKIDYLDKQVAAAFVRRERAMK 455  
Db 394 AMLTFEALSETSEPAKKVVPCKRYNIEPPAPEMWFOOKIDYLDKQVAAAFVRRERAMK 453  
QY 456 EYEFKRYNINLVAKAOKVPEEGMTMOPGTPWPGNNVVDHPGMTQVFLGOSGGLDCNE 515  
Db 454 EYEFKRYNINLVAKAOKVPEEGMTMADGTAMPGNPDHGMQVFLGHSGLDITDNE 513  
QY 516 LPRVYVSREKRGPGYNNHKKAGAMNALVRSALVLTNAPLYLNLCDHYINNSKAIKEAMC 575  
Db 514 LPRVYVSREKRGPGYNNHKKAGAMNALVRSALVLTNGAYLVNDCDHYFNSSKALREAMC 573  
QY 576 FMMDPLGKVCYVQFQREFDIDRHRVYANRVVFFDINMGDLGIGPIYVGTGYCFR 635  
Db 574 FMMDPLGKVCYVQFQREFDIDRHRVYANRVVFFDINMGDLGIGPIYVGTGYCFR 633  
QY 636 ROALYGYDAPKTKRPRSTNCWPKWCCCFGNRKOKKTTPKTEKXLLFFKXENO 695  
Db 634 ROALYGYD-----PVLTEADLEPVYIKSCC--GRRKKKNSYMDQSQR---IMKRTESS 683  
QY 696 SPAYALGELDEAPGAENKAGIYNOQKLEKKFGQSVFVTSLTLENGTLKASAPASL 755  
Db 684 APIFMMEIDIEGIEYEDERSVMSQKLEKRFQSPFIFASTMTQGIPIPSINPASL 743

QY 756 KEAHVISCYEDKTDWKEIGMIGSVTEDIITGFKMHCHGMSIYCI PKRAFKGSAP 815  
 DB 744 KEAHVISCYEDKTDWKEIGMIGSVTEDIITGFKMHCHGMSIYCI PKRAFKGSAP 803  
 QY 816 LMSRLHOUVRLMAGSIEIFPSNCPMTYGGGLKLERPSYNSIYVMTSIPPLAY 875  
 DB 804 INLSRLHOUVRLMAGSIEIFPSNCPMTYGGGLKLERPSYNSIYVMTSIPPLAY 863  
 QY 876 CTLPAICLTGKFTPELNNVSLMFMSLFICIFATSILEMWMSGVGIDDMWRNQPVI 935  
 DB 864 CVLPALICLTGKFTPELNNVSLMFMSLFICIFATSILEMWMSGVGIDDMWRNQPVI 923  
 QY 936 GGVSHLFAVFOGLKVLAVDTSTFTVTSKGD-DEEFSSELYTKMTLLIPTTLILN 994  
 DB 924 GGSASHLFAVFOGLKVLAVDTSTFTVTSKGD-DEEFSSELYTKMTLLIPTTLILN 983  
 QY 995 FIGVAVGSNAINNGYSWGPFLFGKLPFAVVIYVHLYPPLKGLVGRNRTFTVIWISIL 1054  
 DB 984 LVGMVAGISYAINSGYSWGPFLFGKLPFAVVIYVHLYPPLKGLVGRNRTFTVIWISIL 1043  
 QY 1055 LASTFSLMWRIDPFLA-KDDGRLLECGDLC 1085  
 DB 1044 LASTFSLMWRIDPFLA-KDDGRLLECGDLC 1075

## RESULT 14

US-10-160-719-10  
 ; Sequence 10, Application US/10160719  
 ; Publication No. US20030167528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kanwarpal S. Dhugra  
 ; APPLICANT: Timothy G. Helentjaris  
 ; APPLICANT: Benjamin A. Bowen  
 ; APPLICANT: Kun Wang  
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
 ; FILE REFERENCE: 0864  
 ; CURRENT APPLICATION NUMBER: US/10/160,719  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: US/09/371,383A  
 ; PRIOR FILING DATE: 1999-08-06  
 ; PRIOR APPLICATION NUMBER: US 60/096,822  
 ; PRIOR FILING DATE: 1998-08-17  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 10  
 ; LENGTH: 1075  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-10-160-719-10

Query Match 67.4%; Score 3991; DB 12; Length 1075;  
 Best Local Similarity 66.5%; Pred. No. 0;  
 Matches 739; Conservative 143; Mismatches 166; Indels 64; Gaps 15;

QY 1 MEASAGLYAGSHNRELVIYRDDP-PG-PKPRRQNOYQOICGDDVGLAGPFPAC 58  
 DB 1 MAANGMVGAGSHNRELVIYRDDP-PG-PKPRRQNOYQOICGDDVGLAGPFPAC 60  
 QY 59 NECAFPVRCDEYERREKTONCPOCKTRYKRLKGCORVTGDEEDGVDDLDNEFMNDH 118  
 DB 61 NECAFPVRCDEYERREKTONCPOCKTRYKRLKGCORVTGDEEDGVDDLDNEFMNDH 117  
 QY 119 DSGVASEMLYGHMSYGRG-GDPNGAPQAFQOLNPN-VPLITNGQWVD---DIP 166  
 DB 118 DSGVASEMLYGHMSYGRG-GDPNGAPQAFQOLNPN-VPLITNGQWVD---DIP 165  
 QY 167 PEQNALVSPFMGSGKRIHPLRYADPSLPYQPRSMDSKDLAAYGYGVAMKERENMKQ 226  
 DB 166 PDRHSI-----RSPYSSYVPPVPRIVDPKDLNSYLSNVDMWERESNRV 215  
 QY 227 ROER-MHOTGN-----DGGGDDGDADLPLMDBAROOLSRKIPLPSQINPYAMI 275  
 DB 227 ROER-MHOTGN-----DGGGDDGDADLPLMDBAROOLSRKIPLPSQINPYAMI 275

DB 216 KODRMMQVTKNYEABAGDMEGTSGNCE--DMQVVDARLPLSRIVDISNQNLTVRV 273  
 QY 276 IIRLVLTGFEFFHRYVNHVPVADAFALMLISYICIMFMSMILLOFPMPPIEETYLDR 335  
 DB 274 IIRLVLTGFEFFHRYVNHVPVADAFALMLISYICIMFMSMILLOFPMPPIEETYLDR 333  
 QY 336 LSLRFDKGGQSOLAPIDFVSTVDPLKEPPLVTTNTVLSTLSVDYVPDKSCVSDGA 395  
 DB 334 LALVYDRGERSQALPIDVSTVDPLKEPPLVTTNTVLSTLSVDYVPDKSCVSDGA 393  
 QY 396 MLTFEALSETSEFAKRVPPCKRYNI EPRAPVYFOOKIDYLDKQVAMFVRERAMKR 455  
 DB 394 MLTFEALSETSEFAKRVPPCKRYNI EPRAPVYFOOKIDYLDKQVAMFVRERAMKR 453  
 QY 456 EYEFKRYINLVAKAKVPEEGTMDGTPPENNVRDHEGMQVFLGSGGGIDCEGNE 515  
 DB 454 EYEFKRYINLVAKAKVPEEGTMDGTPPENNVRDHEGMQVFLGSGGGIDCEGNE 513  
 QY 516 LPRLVYVSREKRPVYNHKKRAGANMALVRSAVLTNAPYLLNDCDHYINNSKAIKEAMC 575  
 DB 514 LPRLVYVSREKRPVYNHKKRAGANMALVRSAVLTNAPYLLNDCDHYINNSKAIKEAMC 573  
 QY 576 FMNDPPLGKVCYVQFPQRFQDIDRDRYANRVVFPDINNKGIDGIGPITYVGTGCVFR 635  
 DB 574 FMNDPPLGKVCYVQFPQRFQDIDRDRYANRVVFPDINNKGIDGIGPITYVGTGCVFR 633  
 QY 636 RQALYGDAPKTKPSPRTCKMPKCFCCCFENRQKTKTKTEKKLLPFKEBNO 695  
 DB 634 RQALYGDAPKTKPSPRTCKMPKCFCCCFENRQKTKTKTEKKLLPFKEBNO 683  
 QY 696 SPAYALGEIDEAARAGANERKAGIYNOOKLEKKFGSSVFTSTLENGTLKASPSIL 755  
 DB 684 SPAYALGEIDEAARAGANERKAGIYNOOKLEKKFGSSVFTSTLENGTLKASPSIL 743  
 QY 756 KEAHVISCYEDKTDWKEIGMIGSVTEDIITGFKMHCHGMSIYCI PKRAFKGSAP 815  
 DB 744 KEAHVISCYEDKTDWKEIGMIGSVTEDIITGFKMHCHGMSIYCI PKRAFKGSAP 803  
 QY 816 LMSRLHOUVRLMAGSIEIFPSNCPMTYGGGLKLERPSYNSIYVMTSIPPLAY 875  
 DB 804 INLSRLHOUVRLMAGSIEIFPSNCPMTYGGGLKLERPSYNSIYVMTSIPPLAY 863  
 QY 876 CTLPAICLTGKFTPELNNVSLMFMSLFICIFATSILEMWMSGVGIDDMWRNQPVI 935  
 DB 864 CVLPALICLTGKFTPELNNVSLMFMSLFICIFATSILEMWMSGVGIDDMWRNQPVI 923  
 QY 936 GGVSHLFAVFOGLKVLAVDTSTFTVTSKGD-DEEFSSELYTKMTLLIPTTLILN 994  
 DB 924 GGSASHLFAVFOGLKVLAVDTSTFTVTSKGD-DEEFSSELYTKMTLLIPTTLILN 983  
 QY 995 FIGVAVGSNAINNGYSWGPFLFGKLPFAVVIYVHLYPPLKGLVGRNRTFTVIWISIL 1054  
 DB 984 LVGMVAGISYAINSGYSWGPFLFGKLPFAVVIYVHLYPPLKGLVGRNRTFTVIWISIL 1043  
 QY 1055 LASTFSLMWRIDPFLA-KDDGRLLECGDLC 1085  
 DB 1044 LASTFSLMWRIDPFLA-KDDGRLLECGDLC 1075

## RESULT 15

US-10-160-719-34  
 ; Sequence 34, Application US/10160719  
 ; Publication No. US20030167528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kanwarpal S. Dhugra  
 ; APPLICANT: Timothy G. Helentjaris  
 ; APPLICANT: Benjamin A. Bowen  
 ; APPLICANT: Kun Wang  
 ; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
 ; FILE REFERENCE: 0864  
 ; CURRENT APPLICATION NUMBER: US/10/160,719  
 ; CURRENT FILING DATE: 2002-06-03



; PRIOR APPLICATION NUMBER: US/09/371,383A  
 ; PRIOR FILING DATE: 1999-08-06  
 ; PRIOR APPLICATION NUMBER: US 60/096,822  
 ; PRIOR FILING DATE: 1998-08-17  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 34  
 ; LENGTH: 1075  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; US-10-160-719-34

Query Match 67.4%; Score 3991; DB 12; Length 1075;

Best Local Similarity 66.5%; Pred. No. 0; Matches 739; Conservative 143; Mismatches 166; Indels 64; Gaps 15;

QY 1 MEASAGLVAGSHRNRLVYLRDGD-PG-PKPREQNGQVCOICGDVGLAPGADPFVAC 58  
 DB 1 MAANKMGVAGSHRNRLVYLRDGDVPGSAKPTKSANGVCOICGDSVGSATGDVFAVC 60  
 QY 59 NECAFVPCRDYEREREGTQNCPOCKTRYKLGQQRVTGDEEDGVDDLDNEFMWGH 118  
 DB 61 NECAFVPCRDYEREREGTQNCPOCKTRYKLGQQRVTGDEEDGVDDLDNEFMWGH 117  
 QY 119 DSGVASEMLYGMSTYGRG-----GDRNGAPQAFQNLPR--VPLLTNGQMVU---DIP 166  
 DB 118 -----KQSGKGPBEMQLQDDBDLSSSARHEPHRIPLRTSGQISGETPDAS 165  
 QY 167 PEQHALVPSFMGGGKRINPLPYADPSLPVQPSMDPSKOLAAVGYGSVAMKERMENMKQ 226  
 DB 166 PDHST-----RSPTSSIVDPSPVPRVIVDPKDLNYSGLNSVDMKERVESWRY 215  
 QY 227 ROER-MHQTGN-----DGGGDDDDADLPLMDEARQULSRKIPLESSQINPYRMI 275  
 DB 216 KQDKMMQVYTKYPEARAGDMEGTSGNGE--DWQVVDARLPLSRIVPISNQLNYRYV 273  
 QY 276 IIRLVVLCFFHYRYVMHPNDAPALMLISVCEIWPAMSWILDOFPKMPRIERTYLDK 335  
 DB 274 IIRLVVLCFFHYRYVMHPNDAPALMLISVCEIWPAMSWILDOFPKMPRIERTYLDK 333  
 QY 336 LSLRPFKEGQPSQAPIDPFVSTVDPLKEPPLVTNTVLSILSDYFVVDKVSQVSDGGA 395  
 DB 334 LALRYBREGPSQAPIDPFVSTVDPLKEPPLVTNTVLSILSDYFVVDKVSQVSDGGS 393  
 QY 396 AMLTPEALSTSEFAKKWVPFCCKRYNIEPRAPEMYFOQKIDYLKDRVAANFVERRAMKR 455  
 DB 394 AMLTPEALSTSEFAKKWVPFCCKRYNIEPRAPEMYFOQKIDYLKDRVAANFVERRAMKR 453  
 QY 456 EYEEFKRINALVAKAKQVPEBEGTMDGTWPAGNNVRDHPGMIOVFLGHSGLDITDNE 515  
 DB 454 EYEEFKRINALVAKAKQVPEBEGTMDGTWPAGNNVRDHPGMIOVFLGHSGLDITDNE 513  
 QY 516 LPLRVVYSREKRPQGVNHHKAGAMNALLVRSAYLTNAAYLNLDCDHYINNSKAKEAMC 575  
 DB 514 LPLRVVYSREKRPQGVNHHKAGAMNALLVRSAYLTNAAYLNLDCDHYINNSKAKEAMC 573  
 QY 576 FMNDPLGKVCYVQFQREIDRHRDVRANRVVFFDINMKGLDGIQGP1YVGTGCVFR 635  
 DB 574 FMNDPLGKVCYVQFQREIDRHRDVRANRVVFFDINMKGLDGIQGPVYGTGCVFN 633  
 QY 636 RQALYGDAPKTKKPPERTCNCMPKMFCCCCFGNKKOKKTKTKTEKKLLFFKKEENQ 695  
 DB 634 RQALYGD-----PVLTEADLEPNIVIKSC--GRKKKKKSYNDSSR---IMKRTSS 683  
 QY 696 SPYALGEIDEAPGAENKAGIVNOOKLEKKFGQSVFVSTLLENGTLKASPASL 755  
 DB 684 APYPMEDIEEGYEDERSVLSQKLEKRPQSPIFASTFMQGGIPSTNPASL 743  
 QY 756 KEAIIHVISGVEDKTDWKEIGWLYGSVTEDILTGFMHCHGWRSTYCIPIKVAFKGAP 815  
 DB 744 KEAIIHVISGVEDKTDWKEIGWLYGSVTEDILTGFMHCHGWRSTYCIPIKVAFKGAP 803  
 QY 816 LNLSDRLHQLRMAAGSIEIFSNHCPWMYGGGLKFLERFSYINSIVYPMWSIPLAY 875

DB 804 INLSRLNQLVRLMAGSIBILLSRHCP1WYGNRKLLERLATINTIVYITSIPLAY 863  
 QY 876 CTLPAICLLTSGFTPELANVASIMFMSLFCIPATSTILENMGSGVIGIDWNRNEQFVY 935  
 DB 864 CULPAICLLTNNKFIIPRISNAGMFILLFASIFATGILELRWGSVGIEDWNRNEQFVY 923  
 QY 936 GGVSHLFAVFOGLIKVAGVDTSPTVYSKGD--DEESELVTKRMTLLIPTTILLN 994  
 DB 924 GGVSHLFAVFOGLIKVAGVDTSPTVYSKGD--DEESELVTKRMTLLIPTTILLN 983  
 QY 995 FIGVAVGSNAINNGVESMGPLFGKLPFAFVIVVHLYPFLKGLVQRORPTPIVWSIL 1054  
 DB 984 LVGMVAGISYAINSGYQSMGPLFGKLPFAFVIVVHLYPFLKGLVQRORPTPIVWSIL 1043  
 QY 1055 LASIFSLMVRIDPFLA-KDDGPLEEGCLDC 1085  
 DB 1044 LASIFSLMVRIDPFLA-KDDGPLEEGCLDC 1075

Search completed: December 15, 2003, 13:10:07  
 Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:01:50 ; Search time 22 Seconds

(without alignment) 2088.618 Million cell updates/sec

Title: US-09-720-383C-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNNRELVI.....DFLAKDGPILLEGGIDCN 1086

Scoring table:

BLOSUM62

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4250.5	71.8	1084	4	US-09-221-013A-8
2	3945.5	66.7	1081	4	US-09-221-013A-6
3	3941.5	66.5	1081	4	US-09-221-013A-12
4	3876.5	65.5	1065	4	US-09-221-013A-10
5	3552	60.0	881	3	US-08-960-048-8
6	3552	60.0	881	3	US-08-960-048-8
7	3469.5	58.6	974	3	US-08-960-048-6
8	3469.5	58.6	974	3	US-08-960-048-6
9	2902	49.0	685	3	US-08-960-048-7
10	2902	49.0	685	3	US-08-960-048-7
11	2436.5	41.2	529	4	US-09-221-013A-2
12	1950.5	33.0	647	4	US-09-221-013A-14
13	252	4.3	3031	1	US-07-689-008-2
14	241.5	4.1	693	3	US-08-960-048-11
15	241.5	4.1	693	3	US-08-960-048-11
16	237	4.0	756	3	US-08-960-048-10
17	237	4.0	756	3	US-08-960-048-10
18	237	4.0	756	4	US-09-522-474-3
19	237	4.0	756	4	US-09-522-474-3
20	227.5	3.8	861	3	US-08-960-048-12
21	227.5	3.8	861	3	US-08-960-048-12
22	201	3.4	723	3	US-08-960-048-9
23	201	3.4	723	3	US-08-960-048-9
24	201	3.4	723	4	US-08-541-939-2
25	201	3.4	723	4	US-08-541-939-2
26	118.5	2.0	357	1	PCT-US91-01726-4
27	117	2.0	1817	4	US-08-119-773-4
					US-09-004-838-125

28	110	1.9	1912	4	US-08-913-832A-2	Sequence 2, Appli
29	110	1.9	1912	4	US-09-249-181A-2	Sequence 2, Appli
30	110	1.9	1912	4	US-09-158-707-2	Sequence 2, Appli
31	109.5	1.8	357	1	US-08-119-773-6	Sequence 6, Appli
32	108.5	1.8	346	1	US-08-119-773-5	Sequence 5, Appli
33	108.5	1.8	357	1	US-08-119-773-2	Sequence 2, Appli
34	105	1.8	904	4	US-09-252-991A-19257	Sequence 19257, A
35	104	1.8	944	2	US-08-867-941-23	Sequence 23, Appli
36	104	1.8	944	3	US-09-074-658-23	Sequence 23, Appli
37	102	1.7	586	4	US-09-252-991A-28275	Sequence 28275, A
38	101.5	1.7	418	4	US-09-134-001C-4051	Sequence 4051, Ap
39	100	1.7	308	4	US-09-549-848B-37	Sequence 37, Appli
40	99.5	1.7	1416	1	US-08-061-465-4	Sequence 4, Appli
41	98.5	1.7	505	3	US-08-657-868B-4	Sequence 4, Appli
42	98.5	1.7	505	4	US-09-532-180A-4	Sequence 4, Appli
43	98	1.7	351	5	PCT-US91-06418-1	Sequence 1, Appli
44	98	1.7	501	4	US-09-328-352-5227	Sequence 5227, Ap
45	97.5	1.6	996	4	US-09-252-991A-27018	Sequence 27018, A

ALIGNMENTS

RESULT 1									
US-09-221-013A-8									
Sequence 8, Application US/09221013A									
Patent No. 6495740									
GENERAL INFORMATION:									
APPLICANT: Arioli, Antonio									
APPLICANT: Williamson, Richard E.									
APPLICANT: Betzner, Andreas S.									
APPLICANT: Peng, Liangcai									
TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan									
FILE REFERENCE: 96-98									
CURRENT FILING DATE: 1998-12-23									
PRIOR FILING DATE: 1997-06-24									
PRIOR FILING DATE: 1996-06-27									
PRIOR APPLICATION NUMBER: PCT/AU97/00402									
PRIOR FILING DATE: 1997-06-24									
PRIOR APPLICATION NUMBER: AU P00699									
PRIOR FILING DATE: 1996-06-27									
NUMBER OF SEQ ID NOS: 37									
SOFTWARE: Patentin Ver. 2.0									
SEQ ID NO 8									
LENGTH: 1084									
TYPE: PRT									
ORGANISM: Arabidopsis thaliana									
US-09-221-013A-8									
Query Match									
Best Local Similarity 71.8%; Score 4250.5; DB 4; Length 1084;									
Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;									
QY	1	MEASAGLVAGSHNNRELVI	RRDDPGKPREBNGOVCOICGPDVGLAPGDPFVACNE	60					
DB	1	MNTGRLGLASHNNRELVI	LNADSSARIRSVQELSGQTCQCGRIELTVSSSELPVACNE	60					
QY	61	CAFPVCRDGYEEREGTQNC	PQCKTRYKRLKGCQRYTG-DEEBDGVDDLDFNBNMGHD	119					
DB	61	CAPVCRDGYEEREGTQNC	PQCKTRYKRLKGCQRYTG-DEEBDGVDDLDFNBNMGHD	119					
QY	120	SQVAESLTHYHMGVYRG	CGDPNGAPQAFQNLNPNVPLTNGMVDIDPEQALVPSFMGG	179					
DB	120	PEHAHAELSSRLTGTG	SGDAPPP-SSQPLPLTYCDEDDMYSDBRHATLVPSGTG	174					
QY	180	GGRKIHPLPYADPLP	QVPRSMDSKDLAAYGYSVAMKRMENMKORO-ERM-...HOT	234					
DB	175	YGNRYPRFPFDSSAP	POARKSMVPOKDIABEGYSVAMKRMENMKORO-ERM-...HOT	234					
QY	235	GNDGGG-...DDGDADL	PLMDARQQLSRKIPSSQINPYRMIIIRLVVGLFFPHYR	290					
DB	235	GNNRGSDDDDELDD	PPMPPMMDDEGRQPLSRRLPIRSSRINPYRNLICRLALDLPFFHYR	294					
QY	291	VMHVNDAFALMLIV	ICEIFAMSWILDQPPKFPPIERTETYLDRLSLRFDKGQPSOLA	350					

```

Db      295 ILHPVADAYGLMTSLVTCIEMFAVSWILQDFPKMYIEIEETLUDRLSLTEKRGKPSGLA 354
Qy      351 PIDFVSTVDPLKEPPLVTTNTYLSLSVDYVDKXSCYVSDGAAMLTFEALSETSEPA 410
Db      355 PVDVFASTVDPLKEPPLITANTVLSILAVDYVDKACVSNNGAAMLFELSDPADRA 414
Qy      411 KKWVPECKKYNTERPAPWYFOOKIDYLDKVAANVVERBAMKREYEEFKYRINALVAK 470
Db      415 TKWVPECKKFNTERPAPWYFSQMDYLNKXKHPAVRRRRAKMYEEFKYKYNALVAT 474
Qy      471 AOKVPEEGTMDQGTMPGNVVRDHPGMIOVFLGOSGGLDCEGNELPRLVYVSRERPKY 530
Db      475 AOKVPEERTMDQGTMPGNVVRDHPGMIOVFLGHSVGRDTONELPRLVYVSRERPKY 534
Qy      531 NHHKAGANMALVRSAVLTNAPLYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQ 590
Db      535 DHHKAGANNSLVRSAVLSNAPYLLNVDCDHYINNSKAIRESMCFMMDPOSGKVCYVQ 594
Qy      591 PQORFPGIDRHDRYANRVVPDINKGIDGIGPIYVGTGCVPRRQALYGVAPKTKKP 650
Db      595 PQORFPGIDRHDRYANRVVPDINKGIDGIGPIYVGTGCVFRKQALYGVDPAPKTKKP 654
Qy      651 PARTCNWPKWCECCCFGNRKOKT--TKPKTEKKLLFFKKEENQSPAYALGEIDEA- 707
Db      655 PGKTCNWPWKWC-CLCC-GLRKSKTKATDKTNTK-----ETSKOIHLENDGCV 704
Qy      708 -APGAENERKAGIYNQOKLEKKFKGQSVYTSSTLLENGTLKSPASPLKEAIIHVISCY 766
Db      705 IYVSVNEKRSKATOLKLEKKFGQSPVYASAVLQNGVPRNASPACILREAIQVISCY 764
Qy      767 EDKTDWKEIGWYIGSVTEIDLITGFQMHCHGWRSIYICIPKVAFKGAPLNTSDRLHOVL 826
Db      765 OKTEWKEIGWYIGSVTEIDLITGFQMHCHGWRSVYCMKPRAAFKGSAPINISDRLOVL 824
Qy      827 RMAIGSIEIFPSHCHLWYGGGGLKFLRFSYINSIYVPMTSIPLAACLTPAICLLG 886
Db      825 RMAIGSIEIFLSHCHLWYGGGGLKFLRFSYINSIYVPMTSIPLAACLTPAICLLG 884
Qy      887 KETPELNNVASLWMSLFCIFATSILEMWRSVGVGIDDMWRNEQFVIGVSHLFAV 946
Db      885 KETPELNNVASLWMSLFCIFATSILEMWRSVGVGIDDMWRNEQFVIGVSHLFAV 944
Qy      947 OGILKTIAGVDSFTYTSKGGDEERSELYFTKMTLLIPPTLLILNFIIVAGVSNAI 1006
Db      945 OGILKTIAGVDSFTYTSKADGAFSELYIFKMTLLIPPTLLILNFIIVAGVSDAI 1004
Qy      1007 NNGYESNGPLFGKLFPAFVIVYHLVPLKGLVGRORRPTIYVMSIILASTFSLIMVRI 1066
Db      1005 SNGYDSWGPFGKLFPAFVIVYHLVPLKGLMGKODKMPITIVMSIILASTFSLIMVRI 1064
Qy      1067 DPELAKDGPILLECGLDG 1085
Db      1065 NEFVAK-GGFVLEICGLNC 1082

```

RESULT 2  
US-09-221-013A-6

```

; Sequence 6, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221, 013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27

```

```

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-6

Query Match      66.7%; Score 3945.5; DB 4; Length 1081;
Best Local Similarity 66.9%; Pred. No. 0;
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;

Qy      1 MEASAGLVAGSHNRELVIIRDDGPKGPRERONGVYCOICGSDVGLAPGDPPVACNE 60
Db      1 MEASAGLVAGSVRRELRIHESDGTGPKLNNNGQICQICGSDVGLAETGDFVACNE 60
Qy      61 CAFVCRDCEYERREGTONGPOCKTRKXRLKGCORVYGDDEEPDVLDNEPFMDGHS 120
Db      61 CAFVCRDCEYERKDGQCCPCKTRFRNRKGSFVGEDEDDVDVDEENFNY---A 116
Qy      121 QSVASMLYGHMSYRGDPNGAPQAFQNLNPNVPLTNGQWVD---IPRQHALVSPF- 176
Db      117 QGANKA-----RHQRHGEBSFSSSRHSQF-IPLLTHGHTVSGEIRTPDQSVRTTSGP 169
Qy      177 MGGGGRTHPIPIADPSLPVQPRGSDPSKDLAAYGVSVANKEMENKQGR--MHQT 234
Db      170 LGPSDRNAISSPYIDPQPVPRIVDPKDLNSYGLGVNDKEREVEGKQLQEKQMLQMT 229
Qy      235 GN-----DGGDDDDADLPLMDEARQQLSRKIPLESQINPYRMIIIRLVYGF 285
Db      230 KTHGKGEIETGTSNBE--ELQADDTLRPMRNVFIPSSRLPIKRVIIIRLITLIF 287
Qy      286 FEHYRVHPVNDAPALMLISYICBIWFAWSWILQFPKMPRIERTYLDRLSLFDEKQ 345
Db      288 FLOYRTHPVKQAYPLMLTSLVTCIEMFAFSLDQFPWYINNETYLDRLAINDRGE 347
Qy      346 PSQALPIPFYSTVDPLKEPPLVTTNTYLSLSVDYVDKXSCYVSDGAAMLTFEALSE 405
Db      348 PSQALPVVFPYSTVDPLKEPPLVANTVLSLSVDYVDKXACVYSDGSMTLTFEALSE 407
Qy      406 TSEFAKKNVPCCKRNIEPRAPWYFOOKIDYLDKVAANVVERBAMKREYEEFKVAIN 465
Db      408 TAEPAKKNVPCCKRNIEPRAPWYFOOKIDYLDKLOPSVYKERRAMKREYEEFKVAIN 467
Qy      466 ALVAKQVPEEGTMDQGTMPGNVVRDHPGMIOVFLGOSGGLDCEGNELPRLVYVSR 525
Db      468 ALVAKQKIPREGTMDQGTMPGNVVRDHPGMIOVFLGHSGLDTONELPRLVYVSR 527
Qy      526 KRPQNNHKKAGANMALVRSAVLTNAPLYLNLDCDHYINNSKAIKEAMCFMMDPLGK 585
Db      528 KRPQFHHKKAGANMALVRSAVLTNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPALGK 587
Qy      586 VCYVQFQORFPGIDRHDRYANRVVPDINKGIDGIGPIYVGTGCVPRRQALYGVAP 645
Db      588 CCYVQFQORFPGIDRHDRYANRVVPDINKGIDGIGPIYVGTGCVPRRQALYGVAP 647
Qy      646 KTKPEPSRTCNWPKWCECCCFGNRKOKTKTEKKLLFFKKEENQSPAYALGEID 705
Db      648 LTBE-----DLERPIIYKSCGSRKKGSKKXNYEGRGI--NRSSNAPLFMEID 699
Qy      706 EAPGAENERKAGIYNQOKLEKKFGQSVYTSSTLLENGTLKSPASPLKEAIIHVISC 765
Db      700 EGPEGYDDERSIILMSORSVERKGFQSPVFATIMEQGIPTTNPALILKEAIIHVISC 759
Qy      766 YEDKTDWKEIGWYIGSVTEIDLITGFQMHCHGWRSIYICIPKVAFKGAPLNTSDRLHOV 825
Db      760 YEDKTDWKEIGWYIGSVTEIDLITGFQMHAGWSTIYCNPRPAFKGAPLNTSDRLHOV 819
Qy      826 LRMALGSIEIFPSHCHLWYGGGGLKFLRFSYINSIYVPMTSIPLAACLTPAICLLT 885
Db      820 LRMALGSIEILSRCHLWYGGGGLKFLRFLERIAVINTIYITSIPLAACLTPAICLLT 879
Qy      886 GKETPELNNVASLWMSLFCIFATSILEMWRSVGVGIDDMWRNEQFVIGVSHLFAV 945

```

Db 880 DRFIIPISNVAISIMFILLFISIAVTGILELRMSVSIJEDWMRNEQFVIGTSAHLFAV 939  
Qy 946 FQGLLKVIAGVDTSTFTVTSKGD-DEEFSLEYTFKMTLLIPPTLLILNFIQVAVGSN 1004  
Db 940 FQGLLKVIAGIDTFTVTSKATDEDCGPAELIYFKMTALLIPTVLLVNLIGIYAVGSY 999  
Qy 1005 AINNGYEWGFLFGLLFFAFVAVIYHLYPFLKGLVGRKRTPIYIVMSILLASIFSLMV 1064  
Db 1000 AVNSGYOSWGPFLFGLLFFALVIAHLYPFLKGLGRKRTPIYIVMSVLLASIFSLMV 1059  
Qy 1065 RIDPFL 1070  
Db 1060 RINFPV 1065

RESULT 3  
US-09-221-013A-12  
Sequence 12, Application US/09221013A  
Patent No. 6495740  
GENERAL INFORMATION:  
APPLICANT: Arioli, Antonio  
APPLICANT: Williamson, Richard E.  
APPLICANT: Betzner, Andreas S.  
APPLICANT: Peng, Liangcai  
TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
FILE REFERENCE: 96-98  
CURRENT APPLICATION NUMBER: US/09/221.013A  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: PCT/AU97/00402  
PRIOR FILING DATE: 1997-06-24  
PRIOR APPLICATION NUMBER: AU P00699  
PRIOR FILING DATE: 1996-06-27  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12  
LENGTH: 1081  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-221-013A-12

Query Match 66.6%; Score 3941.5; DB 4; Length 1081;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 726; Conservative 140; Mismatches 183; Indels 37; Gaps 11;

Qy 1 MEASAGLVAGSHNRRLVIRRDGPGRPPREGNGVCOJICGDDVGLAPGDDPFAVACNE 60  
Db 1 MEASAGLVAGSHNRRLVIRRDGPGRPPREGNGVCOJICGDDVGLAPGDDPFAVACNE 60  
Qy 61 CAFVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRTVGDDEBDGVDLNEFNMDGHS 120  
Db 61 CAFVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRTVGDDEBDGVDLNEFNMDGHS 120  
Qy 121 QSVAESMLYGHMSYGRGDDPAGAPAFQNLNPNVPLITNGQWVD---IPPEQHALVPSF- 176  
Db 117 QGANAKA-----RQQRHEBEPSSSRHMSQP-IPLLTGHVYSGEIRTPDQSVRTTSGP 169  
Qy 177 MGCGGGRTHPLPYADPSLPVQPRSMPSKDLAAVYGVSVAMEREMNKORER--MQGT 234  
Db 170 LGPSDRNAISSPFIIDPRQPVYRIVDPKDLANSYGLGVNDKKEVKGKLOKEMKMLQMT 229  
Qy 235 GN-----DGGDDDDADLPLMDBAROOLSKRILPSSQINPYEMIIIRLVVLGF 285  
Db 230 GKYHEGKGGEIEGQSGE--ELQWADOTRLPMGRVVPFIPSSRLTPYVVIILRLILCF 287  
Qy 286 FFHFRVWHPVNDAPALMLISYICETWPMASWILDOFPKWPPIERETTYDLRLSLRDKGQ 345  
Db 288 FLOVRTTHPVKNAVPLMLTVICETWPMASWILDOFPKWPPIERETTYDLRLSLRDKGQ 347  
Qy 346 PSOLAPIDFVSTVDPLKEPPLVTNTVLSISVYVPDKVSCVYSDGAMLPFEALSE 405  
Db 348 PSOLVFPVVFVSTVDPLKEPPLVTNTVLSISVYVPDKVSCVYSDGAMLPFEALSE 407

Qy 406 TSEPAKCVPECKRYNIEPRAPEMYFOOKIDYLDKQKAAANFVRBRAMKEEYEFKRYIN 465  
Db 408 TAEPAAKVPCKKFNIEPRAPERYFAOKIDYLDKQKOPSVKERRAMKEEYEFKRYIN 467  
Qy 466 ALVAKQVPEEGTMDGTWPGNNVYRDPGMIOVLGQSGGIDCEGNEFLRVLVYSRE 525  
Db 468 ALVAKQKIPREGTMDGTWPGNNVYRDPGMIOVLGQSGGIDCEGNEFLRVLVYSRE 527  
Qy 526 KRPQYHKKKAGANNALVRSVAVITNAPYLINDCDHYINNSKAIKEMCFMDPILGKK 585  
Db 528 KRPQYHKKKAGANNALVRSVAVITNAPYLINDCDHYINNSKAIKEMCFMDPILGKK 587  
Qy 586 VCVYQPPORPDGIDRHRDYNRNVFEDIMMKGLDGIQSPYVGTGCVFRQALYGDAP 645  
Db 588 CCYVQPPORPDGIDRHRDYNRNVFEDIMMKGLDGIQSPYVGTGCVFRQALYGDAP 647  
Qy 646 KTKAPSPRTGNCWPKMFCFCCCFNRRKOKTKTKPKTEKKLLPFKKEGNSPAVALGEID 705  
Db 648 LTER-----DLENIIVKCCGSRKKGSKKTYNEKRGI--NRSDSAPLPMEDID 699  
Qy 706 EAPGAENKAGIYNQGLEKKFGQSSVFTSTLLENGTLLKSPASLLKEAIVHISCG 765  
Db 700 EGFEYDERSIILMSQSVKRFQSPVFIATFMEQGIPTTNPATLLKEAIVHISCG 759  
Qy 766 YEDTMDKREIGWYGSVTEIDILTFGRKHCHGRSICYIPKRVAFKSSAPLNSDLRHOV 825  
Db 760 YEDTMDKREIGWYGSVTEIDILTFGRKHCHGRSICYIPKRVAFKSSAPLNSDLRHOV 819  
Qy 826 LRNALGSEIFFSNHCMLTYGGGGLKFLERFSYINSIYVPMWISPLAAYCTLPALCILT 885  
Db 820 LRNALGSEIFFSNHCMLTYGGGGLKFLERFSYINSIYVPMWISPLAAYCTLPALCILT 879  
Qy 886 GKFITPBLNVAISLWFSLFCIFPATSILMRMSGVGIDDMWRNEQFVIGTSAHLFAV 945  
Db 880 DRFIIPISNVAISIMFILLFISIAVTGILELRMSVSIJEDWMRNEQFVIGTSAHLFAV 939  
Qy 946 FQGLLKVIAGVDTSTFTVTSKGD-DEEFSLEYTFKMTLLIPPTLLILNFIQVAVGSN 1004  
Db 940 FQGLLKVIAGIDTFTVTSKATDEDCGPAELIYFKMTALLIPTVLLVNLIGIYAVGSY 999  
Qy 1005 AINNGYEWGFLFGLLFFAFVAVIYHLYPFLKGLVGRKRTPIYIVMSILLASIFSLMV 1064  
Db 1000 AVNSGYOSWGPFLFGLLFFALVIAHLYPFLKGLGRKRTPIYIVMSVLLASIFSLMV 1059  
Qy 1065 RIDPFL 1070  
Db 1060 RINFPV 1065

RESULT 4  
US-09-221-013A-10  
Sequence 10, Application US/09221013A  
Patent No. 6495740  
GENERAL INFORMATION:  
APPLICANT: Arioli, Antonio  
APPLICANT: Williamson, Richard E.  
APPLICANT: Betzner, Andreas S.  
APPLICANT: Peng, Liangcai  
TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
FILE REFERENCE: 96-98  
CURRENT APPLICATION NUMBER: US/09/221.013A  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: PCT/AU97/00402  
PRIOR FILING DATE: 1997-06-24  
PRIOR APPLICATION NUMBER: AU P00699  
PRIOR FILING DATE: 1996-06-27  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 1065  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-221-013A-10



Qy	795	CHGRSITCIKRAVAFKSAAPLNSDLRLHOLVRMLSGIETLFFSHNCLMVMYGGGLKFL	854
Db	590	ARGMSIYICMPGRPRPKSAPLNISDRLNQVLRMLGSVEILFFSHCPIMWYGGRLKFL	649
Qy	855	ERFSYINSIVPMTSIPLLAYCTLPALCILTGKFTPELNNVASLMFMSLFICPATSIIL	914
Db	650	ERFVINTTIVPLTSIPLLICVLPALCILTGKFTIPEISNPASIMFISLFISFATGIL	709
Qy	915	EMRSGVGIIDDMWRNEQFMVVGVSHPFAFPGSLIKTLIAGVDTSFFYTSKGD-DEEPS	973
Db	710	EMRSGVGIIDDMWRNEQFMVVGVSHPFAFPGSLIKTLIAGVDTSFKSDSDGYPFA	769
Qy	974	ELIYFKMTLLIIPPTLLLLNFIQVAVGSNAINNGYESWGBLEFKLEFFAFVVIYHLDPF	1033
Db	770	ELIYFKMTLLIIPPTLLIINLVGVAGISVAINSGYSGWBLPFCKLFFAFVVIYHLDPF	829
Qy	1034	LKGLVGRNRPTIYIVWSILLASIFSLIMWRIDPFLAKDDGPILLEGGGLDC	1085
Db	830	LKGLMGRNRPTIYIVVAAILLASIFSLIMWRIDPFTTRVGPDPQTGTGNCI	881

```

RESULT 6
US-09-838-586-8
; Sequence 8, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/0429,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-09-838-586-8

```

Query Match	50.0%	Score 3552	DB 4	Length 881
Best Local Similarity	73.0%	Pred. No. 0		
Matches 651	Conservative 103	Mismatches 108	Indels 30	Gaps 74
Qy	213	GSVAMKEMEMWKKOQER-----MHQINGDGGD-----DGGDADLPLMDEARQ	257	
Db	1	GNVAMKEVVDGKMLKQDKGALPMNTGTSLAISEGGVGDIDASTDYNNEDALLINDETQ	60	
Qy	258	LSRKIRLPSSQINPYRMIIIRLVYLGEFFHHYVHVPYNDAPALWLSVIGEIPWMSI	317	
Db	61	LSRKPLPSSRINPYRNIYLRVYLSLFHNRINPNRYNAPLWLSVIGEIPWLSML	120	
Qy	318	LDQPPKMPPIRETYLIDLSLRFDEGQSPQLAPIDFVSYTDPLAKEPPLVTTNTVLSL	377	
Db	121	IDQPKMPPIRNETLIDRLALRYDGEPSQSLAANDIVSTVDPKKEPPLVYANTVLSIL	180	
Qy	378	SVDPYVDVSCVSDGAAMLTFEALSERSEFAKMPVFCRKYNIERAPEMYFSQKIDY	437	
Db	181	AVDPYVDVSCVSDGAAMLTFDALSETSEFAKMPVYVKYNIERAPEMYFSQKIDY	240	
Qy	438	LKDVAANFVBERRAMKREYEFEFKRINALVAKAKVBEBCGTMQDGTWPNGNNYRDHG	497	
Db	241	LKDKNHREFVDNRAMKREYEFEFKRINGLVAKAKVBEBCGTMQDGTWPNGNNYRDHG	300	
Qy	498	MIQVFLQSGGLDCEGNELPRLVTVSREKRPCTNNHKKAGAMNALVRYSAVLTNAAYLLN	557	
Db	301	MIQVFLHSGGLDCEGNELPRLVTVSREKRPFGHKKAGAMNALVRYSAVLTNQVYLLN	360	

```

QY 558 LDCCHYINNSAIAIEAECFMMDDPLLGKVCVCPQPPQPDGIDRDRDRANRVPPFDIMK 617
Db 361 LDCCHYINNSAIAIEAECFMMDDPLLGKVCVCPQPPQPDGIDRDRDRANRVPPFDIMK 420
QY 618 GLDDIGQPIYVGTGCVFRRQALVYGDAP-KTKKPSPRFCNQMFKCFCCECFGRKKOKXT 676
Db 421 GLDDIGQPIYVGTGCVFRRQALVYGEPPYKOKKGS-----FLSSLCGGRKKKSKS 471
QY 677 TKPTEKREKULFFKKEENSPAYALGEIDEAAPA--ENEKAGIVNOOKLEKKEGOSSVF 734
Db 472 KKKSSDKKK--SNKHVDVASVFLNEDIEBEVEBAGPDEKSLMSQMSLEKRGQSAF 529
QY 735 VTSTLLENGTLKASPAISLKEAIIHVISCGYEDXTQWKEIGMIYGSVTEDIITGFQDH 794
Db 530 VASTLMEGVGPQATPESLKEAIIHVISCGYEDXTQWGTIEIGMIYGSVTEDIITGFQDH 589
QY 795 CHGRRSIYCIPIKRAFKGSAFLNLSDRLHQVLRALASIEIFPSNHCPLMWYGGGLKFL 854
Db 590 ARGGRSIYCMFKRAFKGSAFLNLSDRLHQVLRALASVEILFPRHCPIWYGGGRLEKFL 649
QY 855 ERFSYINSIYPMWISPLLAQCTLPALCILGKFTPELINNVASIMPSLFIPIFATSIL 914
Db 650 ERFAVINTTIIPLTISPLIYCVLPALCILGKFTIPIETISFASIMFISLISIFATIL 709
QY 915 EMRNSGVGIDDMWMEQOFWVIGVYSLSHFAVFOQLKVIAGVDPSFTYTSKGD-DEEFS 973
Db 710 EMRNSGVGIDBMWMEQOFWVIGVYSLSHFAVFOQLKVIAGVDINFTYTSASBEDGCPA 769
QY 974 ELYYFKMTLLIIPPTLLILNFIGVAVGSNAINNGYESWGLPFGKLPFAFWIVHLYPF 1033
Db 770 ELYYFKMTLLIIPPTLLIINLVGVAVGISVAINSGYSQSWGLPFGKLPFAFWIVHLYPF 829
QY 1034 LKGIWGNRRPTIVIVWSLILASIFSLMWRIIDPPLAKDGPILBEGGLDC 1085
Db 830 LKGLMGQNRPTIVIVWAILLASIFSLMWRIIDPFTYVGPPTQCINIC 881

```

```

RESULT 7
US-08-960-048-6
; Sequence 6, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-6

```

	Query Match	Similarity	Score	DB 3;	Length	974;
Bset Local	59.8%;	Pred.	No. 0;			
Matches	638;	Conservative	152;	Mismatches	168;	Indels 109; Gaps 13.
Oy	33	EQNGVQICGDDVDGLAAGDPFPAACNECAPFCVCRDCYEYERREGTONGCPCKRTYKR--	90			
Dd	3	BSGPVCVTCGCEHVGLANNGEFPVACHHCNPFPIKCSCEHYLAKGRKACLGGSTYDENTL	62			
Oy	91	LKGQQRVTGDEEDDGVDLDLDFNEFMWDGHDSVAESMLYGMYSYGRGDPPGAPOAFQLN	150			
Dd	63	LDEVKATGD-----QSTMAALTNKSQD-----	85			
Oy	151	PNVPLLTNGQWVDIIPPEQHALLVPSFNMGCGSKRIHPLEPYADPSPLEVORPRMSDKDLAAV	210			
Dd	86	-----VGIAHRHISVSATLLDSEM-----AE	105			





```

Db      NTIVPFTSLPIAVCSLPALICLLTGKFIIFPLSLNASLFLGFLSLIIVAVLBSNG 807
Qy      748
Db      VQIDMMREOFWVIGVSSH.PAVPOGLIKVIAGVDSFTVTSKGDDSESELYTFKM 980
Qy      921
Db      VSIEDLMREOFWVIGVSAHLFAVFOGFLKMLAGIDTFTVTAKAADADGEGELYIVM 867
Qy      808
Qy      981 TLLLPPTLLLNFINFVAVGSNAINNGESMGLFGLFPFAFWIYHLYPEFLKGLVR 1040
Db      868 TLLLPPTLLLNFINFVAVGSNAINNGESMGLFGLFPFAFWIYHLYPEFLKGLVR 927
Qy      1041 QNRPTVIVMSILASIFSLMVRIDPFLAKDDGPLEE--CGHDC 1085
Db      928 QNRPTVIVMSILASIFSLMVRIDPFLAKDDGPLEE--CGHDC 974

```

```

RESULT 9
US-08-960-048-7
; Sequence 7, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-7

```

```

Query Match      49.0%; Score 2902; DB 3; Length 685;
Best Local Similarity 75.0%; Pred. No. 9.9e-284;
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

Qy      410 AKKVPFCRKYNIIEPRADWYFOQKIDYLDKVAANFVRRRANKREYEEFKVINALVA 469
Db      1 ARWVPFCCKHNVPRADWYFOQKIDYLDKVAANFVRRRANKREYEEFKVINALVA 60
Qy      470 KAQKVPBEGMWQDGTWPGNNVRDHPGMIQVYLGAGALDVGKELPRLVYVSREKPG 529
Db      61 KAQKVPBEGMWQDGTWPGNNVRDHPGMIQVYLGAGALDVGKELPRLVYVSREKPG 120
Qy      530 YNHHKKAAMALYVSAVLTNAPFLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYV 589
Db      121 YNHHKKAAMALYVSAVLTNAPFLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYV 180
Qy      590 QPQRFQIDRDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRQALYGYDAPKTK 649
Db      181 QPQRFQIDRDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRQALYGYDAPKTK 240
Qy      650 PPSRTCNCPKWCFCGCCFCGNRKOKTKTKTEKK-----KLLPFKK-----EENQS 696
Db      241 PPSRTCNCPKWCFCGCCFCGNRKOKTKTKTEKK-----KLLPFKK-----EENQS 296
Qy      697 PAYALGEIDEAPGAEN--EKAGIVNOQKLEKFGQSSVFVSTLLNENGLTKSASPALL 755
Db      297 PAYALGEIDEAPGAEN--EKAGIVNOQKLEKFGQSSVFVSTLLNENGLTKSASPALL 356
Qy      756 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 815
Db      357 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 416
Qy      816 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 875
Db      417 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 476

```

```

Qy      876 CTLPALICLLTGKFIITPELNVASLPMFSLFICIFATSLIEMRSGVGDIDMMRROFVNI 935
Db      477 CTLPALICLLTGKFIITPELNVASLPMFSLFICIFATSLIEMRSGVGDIDMMRROFVNI 536
Qy      936 GGVSAHLPFAVPOGLIKVIAGVDSFTVTSKGDDSESELYTFKMTLLLPPTLLLNIN 995
Db      537 GGVSAHLPFAVPOGLIKVIAGVDSFTVTSKGDDSESELYTFKMTLLLPPTLLLNIN 596
Qy      996 IGVAVGSNAINNGESMGLFGLFPFAFWIYHLYPEFLKGLVORQNRPTVIVMSIL 1055
Db      597 IGVAVGSNAINNGESMGLFGLFPFAFWIYHLYPEFLKGLVORQNRPTVIVMSIL 656
Qy      1056 ASIFSLMVRIDPFLAKDDGPLEE--CGHDC 1084
Db      657 ASIFSLMVRIDPFLAKDDGPLEE--CGHDC 974

```

```

RESULT 10
US-09-838-586-7
; Sequence 7, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-586-7

```

```

Query Match      49.0%; Score 2902; DB 4; Length 685;
Best Local Similarity 75.0%; Pred. No. 9.9e-284;
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

Qy      410 AKKVPFCRKYNIIEPRADWYFOQKIDYLDKVAANFVRRRANKREYEEFKVINALVA 469
Db      1 ARWVPFCCKHNVPRADWYFOQKIDYLDKVAANFVRRRANKREYEEFKVINALVA 60
Qy      470 KAQKVPBEGMWQDGTWPGNNVRDHPGMIQVYLGAGALDVGKELPRLVYVSREKPG 529
Db      61 KAQKVPBEGMWQDGTWPGNNVRDHPGMIQVYLGAGALDVGKELPRLVYVSREKPG 120
Qy      530 YNHHKKAAMALYVSAVLTNAPFLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYV 589
Db      121 YNHHKKAAMALYVSAVLTNAPFLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYV 180
Qy      590 QPQRFQIDRDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRQALYGYDAPKTK 649
Db      181 QPQRFQIDRDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRQALYGYDAPKTK 240
Qy      650 PPSRTCNCPKWCFCGCCFCGNRKOKTKTKTEKK-----KLLPFKK-----EENQS 696
Db      241 PPSRTCNCPKWCFCGCCFCGNRKOKTKTKTEKK-----KLLPFKK-----EENQS 296
Qy      697 PAYALGEIDEAPGAEN--EKAGIVNOQKLEKFGQSSVFVSTLLNENGLTKSASPALL 755
Db      297 PAYALGEIDEAPGAEN--EKAGIVNOQKLEKFGQSSVFVSTLLNENGLTKSASPALL 356
Qy      756 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 815
Db      357 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 416
Qy      816 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 875
Db      417 KEAHIVISCGYEDKTDWGEIKGIMYGSVTEBILTFPKHCHGMSIYICIPKRVAFKGSAP 476

```

```
Db 417 INLSRHLQVLRWALGSVEIFLSRHCPLYVGGGKLTWLERLAYINTIYPTPTSLPLAY 476
Qy 876 CTLPALCLTGTGKFTPELNNVASLWMSLFICIFATSIEMRSGVIGIDMWRNRFQWY 935
Db 477 CTIPAVCLTGTGFTPTLSNLSVWFLAFLSLANGVELRMSGVSIDMWRNRFQWY 536
Qy 936 GGVSSHLPFAVFOGLKLVLAGVDTSFTVTSKGDDSEFSELYTPFKWTTLLIPTTLLILNF 995
Db 537 GGVSAHLFAVFOGLKLVLAGVDTNFTVTAADTEFEGELYLFKWTLLIIPETLLILNM 596
Qy 996 IGVAVGSNAINGVSGSWGPLFGKLFPAFWVYLHLYPFLKGVGRQNTPTIYVWSTLL 1055
Db 597 VGVAVGSNAINGVSGSWGPLFGKLFPAFWVYLHLYPFLKGVGRQNTPTIYVWSTLL 656
Qy 1056 ASIFSLMWYRIDPFLAKDDGPLEECGLD 1084
Db 657 ASIFSLMWYRIDPFLPKGTGPVAKQCGVE 685
```

```
RESULT 11
US-09-221-013A-2
; Sequence 2, Application US/09221013A
; Patent No. 6435740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-2
```

```
Query Match 41.2%; Score 2436.5; DB 4; Length 629;
Best Local Similarity 71.0%; Pred. No.9.3e-237;
Matches 441; Conservative 78; Mismatches 93; Indels 9; Gaps 3;

Qy 451 RANKREYEEFKVRIINALVAKQVPEEGWTMDGTMPGNNVVDHFGMTQVFLGSGGLD 510
Db 1 RANKREYEEFKVRIINALVAKQKIPEEGWTMDGTMPGNNTRDHPGMIQVFLGSHGGLD 60
Qy 511 CENELPRLVYVSRERPGNNHKKAGAMALVYRVAVLTNAYLLNLDCHYINNSKAI 570
Db 61 TGNELPRLIYVSRERPGOHKKGAGAMASIRVAVLTNGAYLLNVDCHYFNNSKAI 120
Qy 571 KEAMCFMMDPILGKTCYVQFQPFQDIBDHRDVARNVFDPINNKGDGIGPITYGT 630
Db 121 KEAMCFMMDPILGKTCYVQFQPFQDIBDHRDVARNVFDPINNKGDGIGPITYGT 180
Qy 631 GCVFRQALYGYDAPKTKRPSRTKNCWPKMFCFCCCFGNRKQKTKTKTEKKLLFFK 690
Db 181 GCVFNQALYGYDAPVLT-----DLEPVIIVKSCSGSRKKGSKKTYEKRGII-N 232
Qy 691 KEENQSPALALGSDIAAAGAEENKAGIVNQCLLEKKFGQSSVYVTSITLLENGTISKAS 750
Db 233 RSNSSNPLFNMEDIDGFEQYDERSILMSQSRVEKRFQSPVFIAATMEQGLIPTTN 292
Qy 751 PASLLKALIVISCGYEDKTDWKEIGMIYGSYTEIDILGFKHKGHMSIYCIPIKRVAF 810
Db 293 PASLLKALIVISCGYEDKTDWKEIGMIYGSYTEIDILGFKHKGHMSIYCIPIKRVAF 352
```

```
Qy 811 KGSAPLNSDLRHLQVLRWALGSISIFPSNHCPLWYGGGLKFLERSYINSIYPTPTSI 870
Db 353 KGSAPLNSDLRHLQVLRWALGSIBLLSRHCPIDWYGGGRRLRLERLAYINTIYPTPTSI 412
Qy 871 PLVACTPLAICLTGTGKFTPELNNVASLWMSLFICIFATSIEMRSGVIGIDMWRN 930
Db 413 PLVACTPLAICLTGTGKFTPELNNVASLWMSLFISIAVTGILKKNVSIEDMWRN 472
Qy 931 QFWIVGVSHLFAVFOGLKLVLAGVDTSFTVTSKG-GDDEFESELYTPFKWTTLLIPTT 989
Db 473 QFWIVGVSHTLFAVFOGLKLVLAGINTNFTVTSKATKNQDFAKLYLFKWTALLIPTT 532
Qy 990 LLLNFIQVAVGSNAINGVSGSWGPLFGKLFPAFWVYLHLYPFLKGVGRQNTPTIY 1049
Db 533 VLVNFIQVAVGSNAINGVSGSWGPLFGKLFPAFWVYLHLYPFLKGVGRQNTPTIY 592
Qy 1050 VMSILLASIFSLMWYRIDPFL 1070
Db 593 VMSVLLASIFSLMWYRIDPFL 613
```

```
RESULT 12
US-09-221-013A-14
; Sequence 14, Application US/09221013A
; Patent No. 6435740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-221-013A-14
```

```
Query Match 33.0%; Score 1950.5; DB 4; Length 547;
Best Local Similarity 64.1%; Pred. No.9.4e-188;
Matches 369; Conservative 66; Mismatches 84; Indels 57; Gaps 8;

Qy 1 MEASAGLVAGSHNRNELVIRRDG--PGPPPEQNGVQICGDDVGLAPGDPFYAC 58
Db 1 MAANAGVAGSNNRNEFMIRPDGAPPPAPKAGSVNGVQICGDTYGVGATGVFVAC 60
Qy 59 NECAFVPRDCYERREGTQNCPOCKTRYKRLKCGQVATGDEEDGVDDLDNERNWGH 118
Db 61 NECAFVPRPCYERREKGNQCCPOCKTRYKHKCPRVQGBEEDVDLDNDF----- 115
Qy 119 DQSVASBMLVGHMSYGRGDPNGAPQ-----APQINPVPLTLTNGQWMD----- 163
Db 116 -----HYHGNGKGBEWOIQROGEDVDLSSSRHQBHRIPLTISGQISGEIP 163
Qy 164 DIPPEQHALVPSFMGGGGRKIHPLPYADPSLPVQPRSDPDKDLAAYGVSAWAKERMEN 223
Db 164 DASPDHRHSI-----RSGTSYVDPVVPVPRIVDPKDXLNSYGINSDVQWERVAS 213
Qy 224 WKQROER-MHQTGN-----DGGDDGDDADLPLMDARQOLSKTILPSSQINPY 272
Db 214 WRNKQDKMMQVANYKYPEARGGDMGTSGNGB--DIQVVDADARPLSRIVIPSNQNLV 271
Qy 273 RMIIIRLVLVIGFFHYRVMVPVNDAFALMTISVCEIWFPMWSWTLDDQFPKMPFERET 332
Db 272 RIVITIRLILIMFFQYRVTYHVRDAVGLMLVSVICETIMLVSLDQFPMKYPINKET 331
```





Oy 1026 VIVHLYPFLKGLVGRQNRPTIVIVWSILASIFSLWV 1064  
DB 482 -----PPT-----EMLTIVVSRWV 496

Search completed: December 15, 2003, 13:05:32  
Job time : 27 secs